

## NOVEL G PROTEIN-COUPLED RECEPTOR

### CROSS-REFERENCE TO RELATED APPLICATION

The present application claims priority to U.S. Provisional Application Serial No. 60/198,600, filed 19 April 2000.

### FIELD OF THE INVENTION

The present invention relates generally to the fields of genetics and cellular and molecular biology. More particularly, the invention relates to novel G protein coupled receptors (nGPCRs), to polynucleotides that encode such novel receptors, to reagents such as antibodies, probes, primers and kits comprising such antibodies, probes, primers related to the same, and to methods which use the novel G protein coupled receptors, polynucleotides, or reagents.

### BACKGROUND

The G protein-coupled receptors (GPCRs) form a vast superfamily of cell surface receptors which are characterized by an amino-terminal extracellular domain, a carboxyl-terminal intracellular domain, and a serpentine structure that passes through the cell membrane seven times. Hence, such receptors are sometimes also referred to as seven transmembrane (7TM) receptors. These seven transmembrane domains define three extracellular loops and three intracellular loops, in addition to the amino- and carboxy- terminal domains. The extracellular portions of the receptor have a role in recognizing and binding one or more extracellular binding partners (*e.g.*, ligands), whereas the intracellular portions have a role in recognizing and communicating with downstream molecules in the signal transduction cascade.

The G protein-coupled receptors bind a variety of ligands including calcium ions, hormones, chemokines, neuropeptides, neurotransmitters, nucleotides, lipids, odorants, and even photons, and are important in the normal (and sometimes the aberrant) function of many cell

types. [See generally Strosberg, *Eur. J. Biochem.* 196:1-10 (1991) and Bohm *et al.*, *Biochem J.* 322:1-18 (1997).] When a specific ligand binds to its corresponding receptor, the ligand typically stimulates the receptor to activate a specific heterotrimeric guanine-nucleotide-binding regulatory protein (G-protein) that is coupled to the intracellular portion of the receptor. The G protein in turn transmits a signal to an effector molecule within the cell, by either stimulating or inhibiting the activity of that effector molecule. These effector molecules include adenylate cyclase, phospholipases and ion channels. Adenylate cyclase and phospholipases are enzymes that are involved in the production of the second messenger molecules cAMP, inositol triphosphate and diacylglycerol. It is through this sequence of events that an extracellular ligand stimuli exerts intracellular changes through a G protein-coupled receptor. Each such receptor has its own characteristic primary structure, expression pattern, ligand-binding profile, and intracellular effector system.

Because of the vital role of G protein-coupled receptors in the communication between cells and their environment, such receptors are attractive targets for therapeutic intervention, for example by activating or antagonizing such receptors. For receptors having a known ligand, the identification of agonists or antagonists may be sought specifically to enhance or inhibit the action of the ligand. Some G protein-coupled receptors have roles in disease pathogenesis (*e.g.*, certain chemokine receptors that act as HIV co-receptors may have a role in AIDS pathogenesis), and are attractive targets for therapeutic intervention even in the absence of knowledge of the natural ligand of the receptor. Other receptors are attractive targets for therapeutic intervention by virtue of their expression pattern in tissues or cell types that are themselves attractive targets for therapeutic intervention. Examples of this latter category of receptors include receptors expressed in immune cells, which can be targeted to either inhibit autoimmune responses or to enhance immune responses to fight pathogens or cancer; and receptors expressed in the brain or other neural organs and tissues, which are likely targets in the treatment of schizophrenia, depression, bipolar disease, or other neurological disorders. This latter category of receptor is also useful as a marker for identifying and/or purifying (*e.g.*, via fluorescence-activated cell

sorting) cellular subtypes that express the receptor. Unfortunately, only a limited number of G protein receptors from the central nervous system (CNS) are known. Thus, a need exists for G protein-coupled receptors that have been identified and show promise as targets for therapeutic intervention in a variety of animals, including humans.

### SUMMARY OF THE INVENTION

The present invention addresses the need identified above in that it provides a gene encoding a heretofore unknown G protein-coupled receptor, an nGPCR termed "Con-218"; constructs and recombinant host cells incorporating the gene; the novel GPCR polypeptides encoded by the gene; antibodies to the polypeptides; kits employing the polynucleotides and polypeptides, and methods of making and using all of the foregoing.

Although the sequences provided are particular human and rat sequences, the invention is intended to include within its scope other human allelic variants; non-human mammalian forms of Con-218, and other vertebrate forms of Con-218. For example, SEQ ID NOs 1 and 2 represent the polynucleotide and polypeptide of a human Con-218, while SEQ ID NOs 3 and 4 represent those of a rat Con-218. When referring specifically to rat Con-218, the terminology "Con-218-RN" will be used.

The present invention relates to an isolated nucleic acid molecule that comprises a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence homologous to a sequence selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:4, or a fragment thereof. The nucleic acid molecule encodes at least a portion of Con-218. In some embodiments, the nucleic acid molecule comprises a sequence that encodes a polypeptide comprising a sequence selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:4, or a fragment thereof. In some embodiments, the nucleic acid molecule comprises a sequence homologous to SEQ ID NO:1 or SEQ ID NO:3, or a fragment thereof. In some embodiments, the nucleic acid molecule comprises the sequence of SEQ ID NO: 1 or 3, and fragments thereof.

According to some embodiments, the present invention provides vectors which comprise the nucleic acid molecule of the invention. In some embodiments, the vector is an expression vector.

According to some embodiments, the present invention provides host cells which comprise the vectors of the invention. In some embodiments, the host cells comprise expression vectors.

The present invention provides an isolated nucleic acid molecule comprising a nucleotide sequence complementary to at least a portion of SEQ ID NO: 1 or 3, said portion comprising at least 10 nucleotides.

The present invention provides a method of producing a polypeptide comprising a sequence selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:4, or a homolog or fragment thereof. The method comprises the steps of introducing a recombinant expression vector that includes a nucleotide sequence that encodes the polypeptide into a compatible host cell, growing the host cell under conditions for expression of the polypeptide and recovering the polypeptide.

The present invention provides an isolated antibody which binds to an epitope on a polypeptide comprising a sequence selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:4, or a homolog or fragment thereof.

The present invention provides an method of inducing an immune response in a mammal against a polypeptide comprising a sequence selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:4, or a homolog or fragment thereof. The method comprises administering to a mammal an amount of the polypeptide sufficient to induce said immune response.

The present invention provides a method for identifying a compound which binds Con-218. The method comprises the steps of: contacting Con-218 with a compound and determining whether the compound binds Con-218.

The present invention provides a method for identifying a compound which binds a nucleic acid molecule encoding Con-218. The method comprises the steps of contacting said

nucleic acid molecule encoding Con-218 with a compound and determining whether said compound binds said nucleic acid molecule.

The present invention provides a method for identifying a compound which modulates the activity of Con-218. The method comprises the steps of contacting Con-218 with a compound and determining whether Con-218 activity has been modulated.

The present invention provides a method of identifying an animal homolog of Con-218. The method comprises the steps screening a nucleic acid database of the animal with SEQ ID NO: 1 or 3, or a portion thereof and determining whether a portion of said library or database is homologous to SEQ ID NO: 1 or 3, or portion thereof.

The present invention provides a method of identifying an animal homolog of Con-218. The method comprises the steps screening a nucleic acid library of the animal with a nucleic acid molecule having a sequence of SEQ ID NO:1 or SEQ ID NO:3, or a portion thereof; and determining whether a portion of said library or database is homologous to said sequence of SEQ ID NO:1 or SEQ ID NO:3, or a portion thereof.

Another aspect of the present invention relates to methods of screening a human subject to diagnose a disorder affecting the brain or genetic predisposition therefor. The methods comprise the steps of assaying nucleic acid of a human subject to determine a presence or an absence of a mutation altering an amino acid sequence, expression, or biological activity of at least one nGPCR that is expressed in the brain. Con-218 comprises an amino acid sequence selected from the group consisting of: SEQ ID NO:2 and SEQ ID NO:4, and allelic variants thereof. A diagnosis of the disorder or predisposition is made from the presence or absence of the mutation. The presence of a mutation altering the amino acid sequence, expression, or biological activity of Con-218 correlates with an increased risk of developing the disorder.

The present invention further relates to methods of screening for a Con-218 hereditary schizophrenia genotype in a human patient. The methods comprise the steps of providing a biological sample comprising nucleic acid from the patient, in which the nucleic acid includes

sequences corresponding to alleles of Con-218. The presence of one or more mutations in the Con-218 allele is indicative of a hereditary schizophrenia genotype.

The present invention provides kits for screening a human subject to diagnose schizophrenia or a genetic predisposition therefor. The kits include an oligonucleotide useful as a probe for identifying polymorphisms in a human Con-218. The oligonucleotide comprises 6-50 nucleotides in a sequence that is identical or complementary to a sequence of a wild type human Con-218 gene sequence or Con-218 coding sequence, except for one sequence difference selected from the group consisting of a nucleotide addition, a nucleotide deletion, or nucleotide substitution. The kit also includes a media packaged with the oligonucleotide. The media contains information for identifying polymorphisms that correlate with schizophrenia or a genetic predisposition therefor, the polymorphisms being identifiable using the oligonucleotide as a probe.

The present invention further relates to methods of identifying Con-218 allelic variants that correlates with mental disorders. The methods comprise the steps of providing biological samples that comprise nucleic acid from a human patient diagnosed with a mental disorder, or from the patient's genetic progenitors or progeny, and detecting in the nucleic acid the presence of one or more mutations in an nGPCR that is expressed in the brain. Con-218 comprises an amino acid sequence selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:4, and allelic variants thereof. The nucleic acid includes sequences corresponding to the gene or genes encoding Con-218. The one or more mutations detected indicate an allelic variant that correlates with a mental disorder.

The present invention further relates to purified polynucleotides comprising nucleotide sequences encoding alleles of Con-218 from a human with schizophrenia. The polynucleotide hybridizes to the complement of SEQ ID NO:1 or SEQ ID NO:3 under the following hybridization conditions: (a) hybridization for 16 hours at 42°C in a hybridization solution comprising 50% formamide, 1% SDS, 1 M NaCl, 10% dextran sulfate and (b) washing 2 times for 30 minutes at 60°C in a wash solution comprising 0.1x SSC and 1% SDS. The

polynucleotide that encodes Con-218 amino acid sequence of the human differs from SEQ ID NO:2 or SEQ ID NO:4 by at least one residue.

The present invention also provides methods for identifying a modulator of biological activity of Con-218 comprising the steps of contacting a cell that expresses Con-218 in the presence and in the absence of a putative modulator compound and measuring Con-218 biological activity in the cell. The decreased or increased Con-218 biological activity in the presence versus absence of the putative modulator is indicative of a modulator of biological activity.

The present invention further provides methods to identify compounds useful for the treatment of schizophrenia. The methods comprise the steps of contacting a composition comprising Con-218 with a compound suspected of binding Con-218. The binding between Con-218 and the compound suspected of binding Con-218 is detected. Compounds identified as binding Con-218 are candidate compounds useful for the treatment of schizophrenia.

The present invention further provides methods for identifying a compound useful as a modulator of binding between Con-218 and a binding partner of Con-218. The methods comprise the steps of contacting the binding partner and a composition comprising Con-218 in the presence and in the absence of a putative modulator compound and detecting binding between the binding partner and Con-218. Decreased or increased binding between the binding partner and Con-218 in the presence of the putative modulator, as compared to binding in the absence of the putative modulator is indicative a modulator compound useful for the treatment of schizophrenia.

Another aspect of the present invention relates to methods of purifying a G protein from a sample containing a G protein. The methods comprise the steps of contacting the sample with Con-218 for a time sufficient to allow the G protein to form a complex with Con-218; isolating the complex from remaining components of the sample; maintaining the complex under conditions which result in dissociation of the G protein from Con-218; and isolating said G protein from Con-218.

## DETAILED DESCRIPTION OF THE INVENTION

### Definitions

Various definitions are made throughout this document. Most words have the meaning that would be attributed to those words by one skilled in the art. Words specifically defined either below or elsewhere in this document have the meaning provided in the context of the present invention as a whole and as are typically understood by those skilled in the art.

"Synthesized" as used herein and understood in the art, refers to polynucleotides produced by purely chemical, as opposed to enzymatic, methods. "Wholly" synthesized DNA sequences are therefore produced entirely by chemical means, and "partially" synthesized DNAs embrace those wherein only portions of the resulting DNA were produced by chemical means.

By the term "region" is meant a physically contiguous portion of the primary structure of a biomolecule. In the case of proteins, a region is defined by a contiguous portion of the amino acid sequence of that protein.

The term "domain" is herein defined as referring to a structural part of a biomolecule that contributes to a known or suspected function of the biomolecule. Domains may be co-extensive with regions or portions thereof; domains may also incorporate a portion of a biomolecule that is distinct from a particular region, in addition to all or part of that region. Examples of GPCR protein domains include, but are not limited to, the extracellular (*i.e.*, N-terminal), transmembrane and cytoplasmic (*i.e.*, C-terminal) domains, which are co-extensive with like-named regions of GPCRs; each of the seven transmembrane segments of a GPCR; and each of the loop segments (both extracellular and intracellular loops) connecting adjacent transmembrane segments.

As used herein, the term "activity" refers to a variety of measurable indicia suggesting or revealing binding, either direct or indirect; affecting a response, *i.e.* having a measurable affect in response to some exposure or stimulus, including, for example, the affinity of a compound for



directly binding a polypeptide or polynucleotide of the invention, or, for example, measurement of amounts of upstream or downstream proteins or other similar functions after some stimulus or event.

Unless indicated otherwise, as used herein, the abbreviation in lower case (gpcr) refers to a gene, cDNA, RNA or nucleic acid sequence, while the upper case version (GPCR) refers to a protein, polypeptide, peptide, oligopeptide, or amino acid sequence.

As used herein, the term “antibody” is meant to refer to complete, intact antibodies, and Fab, Fab’, F(ab)2, and other fragments thereof. Complete, intact antibodies include monoclonal antibodies such as murine monoclonal antibodies, chimeric antibodies and humanized antibodies.

As used herein, the term “binding” means the physical or chemical interaction between two proteins or compounds or associated proteins or compounds or combinations thereof. Binding includes ionic, non-ionic, Hydrogen bonds, Van der Waals, hydrophobic interactions, etc. The physical interaction, the binding, can be either direct or indirect, indirect being through or due to the effects of another protein or compound. Direct binding refers to interactions that do not take place through or due to the effect of another protein or compound but instead are without other substantial chemical intermediates. Binding may be detected in many different manners. As a non-limiting example, the physical binding interaction between a Con-218 of the invention and a compound can be detected using a labeled compound. Alternatively, functional evidence of binding can be detected using, for example, a cell transfected with and expressing a Con-218 of the invention. Binding of the transfected cell to a ligand of the nGPCR that was transfected into the cell provides functional evidence of binding. Other methods of detecting binding are well-known to those of skill in the art.

As used herein, the term “compound” means any identifiable chemical or molecule, including, but not limited to, small molecule, peptide, protein, sugar, nucleotide, or nucleic acid, and such compound can be natural or synthetic.

As used herein, the term “complementary” refers to Watson-Crick basepairing between nucleotide units of a nucleic acid molecule.

As used herein, the term “contacting” means bringing together, either directly or indirectly, a compound into physical proximity to a polypeptide or polynucleotide of the invention. The polypeptide or polynucleotide can be in any number of buffers, salts, solutions etc. Contacting includes, for example, placing the compound into a beaker, microtiter plate, cell culture flask, or a microarray, such as a gene chip, or the like, which contains the nucleic acid molecule, or polypeptide encoding the nGPCR or fragment thereof.

As used herein, the phrase “homologous nucleotide sequence,” or “homologous amino acid sequence,” or variations thereof, refers to sequences characterized by a homology, at the nucleotide level or amino acid level, of at least the specified percentage. Homologous nucleotide sequences include those sequences coding for isoforms of proteins. Such isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA. Alternatively, isoforms can be encoded by different genes. Homologous nucleotide sequences include nucleotide sequences encoding for a protein of a species other than humans, including, but not limited to, mammals. Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. A homologous nucleotide sequence does not, however, include the nucleotide sequence encoding other known GPCRs. Homologous amino acid sequences include those amino acid sequences which contain conservative amino acid substitutions and which polypeptides have the same binding and/or activity. A homologous amino acid sequence does not, however, include the amino acid sequence encoding other known GPCRs. Percent homology can be determined by, for example, the Gap program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, Madison WI), using the default settings, which uses the algorithm of Smith and Waterman (Adv. Appl. Math., 1981, 2, 482-489, which is incorporated herein by reference in its entirety).

As used herein, the term "isolated" nucleic acid molecule refers to a nucleic acid molecule (DNA or RNA) that has been removed from its native environment. Examples of isolated nucleic acid molecules include, but are not limited to, recombinant DNA molecules

contained in a vector, recombinant DNA molecules maintained in a heterologous host cell, partially or substantially purified nucleic acid molecules, and synthetic DNA or RNA molecules.

As used herein, the terms “modulates” or “modifies” means an increase or decrease in the amount, quality, or effect of a particular activity or protein.

As used herein, the term “oligonucleotide” refers to a series of linked nucleotide residues which has a sufficient number of bases to be used in a polymerase chain reaction (PCR). This short sequence is based on (or designed from) a genomic or cDNA sequence and is used to amplify, confirm, or reveal the presence of an identical, similar or complementary DNA or RNA in a particular cell or tissue. Oligonucleotides comprise portions of a DNA sequence having at least about 10 nucleotides and as many as about 50 nucleotides, preferably about 15 to 30 nucleotides. They are chemically synthesized and may be used as probes.

As used herein, the term “probe” refers to nucleic acid sequences of variable length, preferably between at least about 10 and as many as about 6,000 nucleotides, depending on use. They are used in the detection of identical, similar, or complementary nucleic acid sequences. Longer length probes are usually obtained from a natural or recombinant source, are highly specific and much slower to hybridize than oligomers. They may be single- or double-stranded and carefully designed to have specificity in PCR, hybridization membrane-based, or ELISA-like technologies.

The term “preventing” refers to decreasing the probability that an organism contracts or develops an abnormal condition.

The term “treating” refers to having a therapeutic effect and at least partially alleviating or abrogating an abnormal condition in the organism.

The term “therapeutic effect” refers to the inhibition or activation factors causing or contributing to the abnormal condition. A therapeutic effect relieves to some extent one or more of the symptoms of the abnormal condition. In reference to the treatment of abnormal conditions, a therapeutic effect can refer to one or more of the following: (a) an increase in the proliferation, growth, and/or differentiation of cells; (b) inhibition (*i.e.*, slowing or stopping) of

cell death; (c) inhibition of degeneration; (d) relieving to some extent one or more of the symptoms associated with the abnormal condition; and (e) enhancing the function of the affected population of cells. Compounds demonstrating efficacy against abnormal conditions can be identified as described herein.

The term “abnormal condition” refers to a function in the cells or tissues of an organism that deviates from their normal functions in that organism. An abnormal condition can relate to cell proliferation, cell differentiation, cell signaling, or cell survival. An abnormal condition may also include obesity, diabetic complications such as retinal degeneration, and irregularities in glucose uptake and metabolism, and fatty acid uptake and metabolism.

Abnormal cell proliferative conditions include cancers such as fibrotic and mesangial disorders, abnormal angiogenesis and vasculogenesis, wound healing, psoriasis, diabetes mellitus, and inflammation.

Abnormal differentiation conditions include, but are not limited to, neurodegenerative disorders, slow wound healing rates, and slow tissue grafting healing rates. Abnormal cell signaling conditions include, but are not limited to, psychiatric disorders involving excess neurotransmitter activity.

Abnormal cell survival conditions may also relate to conditions in which programmed cell death (apoptosis) pathways are activated or abrogated. A number of protein kinases are associated with the apoptosis pathways. Aberrations in the function of any one of the protein kinases could lead to cell immortality or premature cell death.

The term “administering” relates to a method of incorporating a compound into cells or tissues of an organism. The abnormal condition can be prevented or treated when the cells or tissues of the organism exist within the organism or outside of the organism. Cells existing outside the organism can be maintained or grown in cell culture dishes. For cells harbored within the organism, many techniques exist in the art to administer compounds, including (but not limited to) oral, parenteral, dermal, injection, and aerosol applications. For cells outside of

the organism, multiple techniques exist in the art to administer the compounds, including (but not limited to) cell microinjection techniques, transformation techniques and carrier techniques.

The abnormal condition can also be prevented or treated by administering a compound to a group of cells having an aberration in a signal transduction pathway to an organism. The effect of administering a compound on organism function can then be monitored. The organism is preferably a mouse, rat, rabbit, guinea pig or goat, more preferably a monkey or ape, and most preferably a human.

By "amplification" it is meant increased numbers of DNA or RNA in a cell compared with normal cells. "Amplification" as it refers to RNA can be the detectable presence of RNA in cells, since in some normal cells there is no basal expression of RNA. In other normal cells, a basal level of expression exists, therefore in these cases amplification is the detection of at least 1 to 2-fold, and preferably more, compared to the basal level.

As used herein, the phrase "stringent hybridization conditions" or "stringent conditions" refers to conditions under which a probe, primer, or oligonucleotide will hybridize to its target sequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium.

Since the target sequences are generally present in excess, at  $T_m$ , 50% of the probes are occupied at equilibrium. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes, primers or oligonucleotides (e.g. 10 to 50 nucleotides) and at least about 60°C for longer probes, primers or oligonucleotides.

Stringent conditions may also be achieved with the addition of destabilizing agents, such as formamide.

The amino acid sequences are presented in the amino to carboxy direction, from left to right. The amino and carboxy groups are not presented in the sequence. The nucleotide sequences are presented by single strand only, in the 5' to 3' direction, from left to right. Nucleotides and amino acids are represented in the manner recommended by the IUPAC-IUB Biochemical Nomenclature Commission or (for amino acids) by three letters code.

### Polynucleotides

The present invention provides purified and isolated polynucleotides (*e.g.*, DNA sequences and RNA transcripts, both sense and complementary antisense strands, both single- and double-stranded, including splice variants thereof) that encode unknown G protein-coupled receptors heretofore termed Con-218.

The invention provides purified and isolated polynucleotides (*e.g.*, cDNA, genomic DNA, synthetic DNA, RNA, or combinations thereof, whether single- or double-stranded) that comprise a nucleotide sequence encoding the amino acid sequence of the polypeptides of the invention. Such polynucleotides are useful for recombinantly expressing the receptor and also for detecting expression of the receptor in cells (*e.g.*, using Northern hybridization and *in situ* hybridization assays). Such polynucleotides also are useful in the design of antisense and other molecules for the suppression of the expression of Con-218 in a cultured cell, a tissue, or an animal; for therapeutic purposes; or to provide a model for diseases or conditions characterized by aberrant Con-218 expression. Specifically excluded from the definition of polynucleotides of the invention are entire isolated, non-recombinant native chromosomes of host cells. A preferred polynucleotide has a sequence of SEQ ID NO:1 or SEQ ID NO:3, which corresponds to naturally occurring Con-218 sequences. It will be appreciated that numerous other polynucleotide sequences exist that also encode Con-218 having the sequence set forth in sequences selected from the group of sequences consisting of SEQ ID NO: 2 and SEQ ID NO:4, due to the well-known degeneracy of the universal genetic code.

The invention also provides a purified and isolated polynucleotide comprising a nucleotide sequence that encodes a mammalian polypeptide, wherein the polynucleotide hybridizes to a polynucleotide having a sequence of SEQ ID NO: 1 or 3, or the non-coding strand complementary thereto, under the following hybridization conditions:

(a) hybridization for 16 hours at 42°C in a hybridization solution comprising 50% formamide, 1% SDS, 1 M NaCl, 10% dextran sulfate; and

(b) washing 2 times for 30 minutes each at 60°C in a wash solution comprising 0.1% SSC, 1% SDS. Polynucleotides that encode a human allelic variant are highly preferred.

The present invention relates to molecules which comprise the gene sequences that encode Con-218; constructs and recombinant host cells incorporating the gene sequences; the novel GPCR polypeptides encoded by the gene sequences; antibodies to the polypeptides and homologs; kits employing the polynucleotides and polypeptides, and methods of making and using all of the foregoing. In addition, the present invention relates to homologs of the gene sequences and of the polypeptides and methods of making and using the same.

Genomic DNA of the invention comprises the protein-coding region for a polypeptide of the invention and is also intended to include allelic variants thereof. It is widely understood that, for many genes, genomic DNA is transcribed into RNA transcripts that undergo one or more splicing events wherein intron (*i.e.*, non-coding regions) of the transcripts are removed, or “spliced out.” RNA transcripts that can be spliced by alternative mechanisms, and therefore be subject to removal of different RNA sequences but still encode a Con-218 polypeptide, are referred to in the art as splice variants which are embraced by the invention. Splice variants comprehended by the invention therefore are encoded by the same original genomic DNA sequences but arise from distinct mRNA transcripts. Allelic variants are modified forms of a wild-type gene sequence, the modification resulting from recombination during chromosomal segregation or exposure to conditions which give rise to genetic mutation. Allelic variants, like

wild type genes, are naturally occurring sequences (as opposed to non-naturally occurring variants which arise from *in vitro* manipulation).

The invention also comprehends cDNA that is obtained through reverse transcription of an RNA polynucleotide encoding Con-218 (conventionally followed by second strand synthesis of a complementary strand to provide a double-stranded DNA).

A preferred DNA sequence encoding a human Con-218 polypeptide is set out in SEQ ID NO: 1, while a preferred DNA sequence encoding a rat Con-218 polypeptide is set out in SEQ ID NO: 3. A preferred DNA of the invention comprises a double stranded molecule along with the complementary molecule (the “non-coding strand” or “complement”) having a sequence unambiguously deducible from the coding strand according to Watson-Crick base-pairing rules for DNA. Also preferred are other polynucleotides encoding the Con-218 polypeptides of SEQ ID NOs: 2 and 4, which differ in sequence from the polynucleotides of SEQ ID NOs: 1 and 3 by virtue of the well-known degeneracy of the universal nuclear genetic code.

The invention further embraces other species, preferably mammalian, homologs of the human Con-218 DNA. Species homologs, sometimes referred to as “orthologs,” in general, share at least 35%, at least 40%, at least 45%, at least 50%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% homology with human DNA of the invention. Generally, percent sequence “homology” with respect to polynucleotides of the invention may be calculated as the percentage of nucleotide bases in the candidate sequence that are identical to nucleotides in the Con-218 sequences set forth in any of SEQ ID NOs: 1 and 3, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity.

Polynucleotides of the invention permit identification and isolation of polynucleotides encoding related Con-218 polypeptides, such as human allelic variants and species homologs, by well-known techniques including Southern and/or Northern hybridization, and polymerase chain reaction (PCR). Examples of related polynucleotides include human and non-human genomic sequences, including allelic variants, as well as polynucleotides encoding polypeptides



homologous to Con-218 and structurally related polypeptides sharing one or more biological, immunological, and/or physical properties of Con-218. Non-human species genes encoding proteins homologous to Con-218 can also be identified by Southern and/or PCR analysis and are useful in animal models for Con-218 disorders. Knowledge of the sequence of a human Con-218 DNA also makes possible through use of Southern hybridization or polymerase chain reaction (PCR) the identification of genomic DNA sequences encoding Con-218 expression control regulatory sequences such as promoters, operators, enhancers, repressors, and the like. Polynucleotides of the invention are also useful in hybridization assays to detect the capacity of cells to express Con-218. Polynucleotides of the invention may also provide a basis for diagnostic methods useful for identifying a genetic alteration(s) in a Con-218 locus that underlies a disease state or states, which information is useful both for diagnosis and for selection of therapeutic strategies.

According to the present invention, the Con-218 nucleotide sequences disclosed herein may be used to identify homologs of the Con-218, in other animals, including but not limited to humans and other mammals, and invertebrates. Any of the nucleotide sequences disclosed herein, or any portion thereof, can be used, for example, as probes to screen databases or nucleic acid libraries, such as, for example, genomic or cDNA libraries, to identify homologs, using screening procedures well known to those skilled in the art. Accordingly, homologs having at least 50%, more preferably at least 60%, more preferably at least 70%, more preferably at least 80%, more preferably at least 90%, more preferably at least 95%, and most preferably at least 100% homology with Con-218 sequences can be identified.

The disclosure herein of a full-length polynucleotide encoding a Con-218 polypeptide makes readily available to the worker of ordinary skill in the art every possible fragment of the full-length polynucleotide.

One preferred embodiment of the present invention provides an isolated nucleic acid molecule comprising a sequence homologous to SEQ ID NO:1 or SEQ ID NO:3, and fragments

thereof. Another preferred embodiment provides an isolated nucleic acid molecule comprising a sequence of SEQ ID NO:1 or SEQ ID NO:3, and fragments thereof.

As used in the present invention, fragments of Con-218-encoding polynucleotides comprise at least 10, and preferably at least 12, 14, 16, 18, 20, 25, 50, or 75 consecutive nucleotides of a polynucleotide encoding Con-218. Preferably, fragment polynucleotides of the invention comprise sequences unique to the Con-218-encoding polynucleotide sequence, and therefore hybridize under highly stringent or moderately stringent conditions only (*i.e.*, “specifically”) to polynucleotides encoding Con-218 (or fragments thereof). Polynucleotide fragments of genomic sequences of the invention comprise not only sequences unique to the coding region, but also include fragments of the full-length sequence derived from introns, regulatory regions, and/or other non-translated sequences. Sequences unique to polynucleotides of the invention are recognizable through sequence comparison to other known polynucleotides, and can be identified through use of alignment programs routinely utilized in the art, *e.g.*, those made available in public sequence databases. Such sequences also are recognizable from Southern hybridization analyses to determine the number of fragments of genomic DNA to which a polynucleotide will hybridize. Polynucleotides of the invention can be labeled in a manner that permits their detection, including radioactive, fluorescent, and enzymatic labeling.

Fragment polynucleotides are particularly useful as probes for detection of full-length or fragment of Con-218 polynucleotides. One or more polynucleotides can be included in kits that are used to detect the presence of a polynucleotide encoding Con-218, or used to detect variations in a polynucleotide sequence encoding Con-218.

The invention also embraces DNAs encoding Con-218 polypeptides that hybridize under moderately stringent or high stringency conditions to the non-coding strand, or complement, of the polynucleotide in any of SEQ ID NOs: 1 and 3.

Exemplary highly stringent hybridization conditions are as follows: hybridization at 42°C in a hybridization solution comprising 50% formamide, 1% SDS, 1 M NaCl, 10% Dextran

sulfate, and washing twice for 30 minutes at 60°C in a wash solution comprising 0.1 X SSC and 1% SDS. It is understood in the art that conditions of equivalent stringency can be achieved through variation of temperature and buffer, or salt concentration as described Ausubel, *et al.* (Eds.), Protocols in Molecular Biology, John Wiley & Sons (1994), pp. 6.0.3 to 6.4.10.

Modifications in hybridization conditions can be empirically determined or precisely calculated based on the length and the percentage of guanosine/cytosine (GC) base pairing of the probe. The hybridization conditions can be calculated as described in Sambrook, *et al.*, (Eds.), Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York (1989), pp. 9.47 to 9.51.

With the knowledge of the nucleotide sequence information disclosed in the present invention, one skilled in the art can identify and obtain nucleotide sequences which encode Con-218 from different sources (*i.e.*, different tissues or different organisms) through a variety of means well known to the skilled artisan and as disclosed by, for example, Sambrook *et al.*, "Molecular cloning: a laboratory manual", Second Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), which is incorporated herein by reference in its entirety.

For example, DNA that encodes Con-218 may be obtained by screening of mRNA, cDNA, or genomic DNA with oligonucleotide probes generated from the Con-218 gene sequence information provided herein. Probes may be labeled with a detectable group, such as a fluorescent group, a radioactive atom or a chemiluminescent group in accordance with procedures known to the skilled artisan and used in conventional hybridization assays, as described by, for example, Sambrook *et al.*

A nucleic acid molecule comprising any of the Con-218 nucleotide sequences described above can alternatively be synthesized by use of the polymerase chain reaction (PCR) procedure, with the PCR oligonucleotide primers produced from the nucleotide sequences provided herein. See U.S. Patent Numbers 4,683,195 to Mullis *et al.* and 4,683,202 to Mullis. The PCR reaction provides a method for selectively increasing the concentration of a particular nucleic acid sequence even when that sequence has not been previously purified and is present only in a

single copy in a particular sample. The method can be used to amplify either single- or double-stranded DNA. The essence of the method involves the use of two oligonucleotide probes to serve as primers for the template-dependent, polymerase mediated replication of a desired nucleic acid molecule.

A wide variety of alternative cloning and *in vitro* amplification methodologies are well known to those skilled in the art. Examples of these techniques are found in, for example, Berger *et al.*, *Guide to Molecular Cloning Techniques*, Methods in Enzymology 152, Academic Press, Inc., San Diego, CA (Berger), which is incorporated herein by reference in its entirety.

Automated sequencing methods can be used to obtain or verify the nucleotide sequence of Con-218. The Con-218 nucleotide sequences of the present invention are believed to be 100% accurate. However, as is known in the art, nucleotide sequence obtained by automated methods may contain some errors. Nucleotide sequences determined by automation are typically at least about 90%, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of a given nucleic acid molecule. The actual sequence may be more precisely determined using manual sequencing methods, which are well known in the art. An error in a sequence which results in an insertion or deletion of one or more nucleotides may result in a frame shift in translation such that the predicted amino acid sequence will differ from that which would be predicted from the actual nucleotide sequence of the nucleic acid molecule, starting at the point of the mutation.

The nucleic acid molecules of the present invention, and fragments derived therefrom, are useful for screening for restriction fragment length polymorphism (RFLP) associated with certain disorders, as well as for genetic mapping.

The polynucleotide sequence information provided by the invention makes possible large-scale expression of the encoded polypeptide by techniques well known and routinely practiced in the art.

#### Vectors

Another aspect of the present invention is directed to vectors, or recombinant expression vectors, comprising any of the nucleic acid molecules described above. Vectors are used herein either to amplify DNA or RNA encoding Con-218 and/or to express DNA which encodes Con-218. Preferred vectors include, but are not limited to, plasmids, phages, cosmids, episomes, viral particles or viruses, and integratable DNA fragments (*i.e.*, fragments integratable into the host genome by homologous recombination). Preferred viral particles include, but are not limited to, adenoviruses, baculoviruses, parvoviruses, herpesviruses, poxviruses, adeno-associated viruses, Semliki Forest viruses, vaccinia viruses, and retroviruses. Preferred expression vectors include, but are not limited to, pcDNA3 (Invitrogen) and pSVL (Pharmacia Biotech). Other expression vectors include, but are not limited to, pSPORT™ vectors, pGEM™ vectors (Promega), pPROEXvectors™ (LTI, Bethesda, MD), Bluescript™ vectors (Stratagene), pQE™ vectors (Qiagen), pSE420™ (Invitrogen), and pYES2™ (Invitrogen).

Expression constructs preferably comprise Con-218-encoding polynucleotides operatively linked to an endogenous or exogenous expression control DNA sequence and a transcription terminator. Expression control DNA sequences include promoters, enhancers, operators, and regulatory element binding sites generally, and are typically selected based on the expression systems in which the expression construct is to be utilized. Preferred promoter and enhancer sequences are generally selected for the ability to increase gene expression, while operator sequences are generally selected for the ability to regulate gene expression. Expression constructs of the invention may also include sequences encoding one or more selectable markers that permit identification of host cells bearing the construct. Expression constructs may also include sequences that facilitate, and preferably promote, homologous recombination in a host cell. Preferred constructs of the invention also include sequences necessary for replication in a host cell.

Expression constructs are preferably utilized for production of an encoded protein, but may also be utilized simply to amplify a Con-218-encoding polynucleotide sequence. In preferred embodiments, the vector is an expression vector wherein the polynucleotide of the

invention is operatively linked to a polynucleotide comprising an expression control sequence. Autonomously replicating recombinant expression constructs such as plasmid and viral DNA vectors incorporating polynucleotides of the invention are also provided. Preferred expression vectors are replicable DNA constructs in which a DNA sequence encoding Con-218 is operably linked or connected to suitable control sequences capable of effecting the expression of the Con-218 in a suitable host. DNA regions are operably linked or connected when they are functionally related to each other. For example, a promoter is operably linked or connected to a coding sequence if it controls the transcription of the sequence. Amplification vectors do not require expression control domains, but rather need only the ability to replicate in a host, usually conferred by an origin of replication, and a selection gene to facilitate recognition of transformants. The need for control sequences in the expression vector will vary depending upon the host selected and the transformation method chosen. Generally, control sequences include a transcriptional promoter, an optional operator sequence to control transcription, a sequence encoding suitable mRNA ribosomal binding and sequences which control the termination of transcription and translation.

Preferred vectors preferably contain a promoter that is recognized by the host organism. The promoter sequences of the present invention may be prokaryotic, eukaryotic or viral. Examples of suitable prokaryotic sequences include the  $P_R$  and  $P_L$  promoters of bacteriophage lambda (The bacteriophage Lambda, Hershey, A. D., Ed., Cold Spring Harbor Press, Cold Spring Harbor, NY (1973), which is incorporated herein by reference in its entirety; Lambda II, Hendrix, R. W., Ed., Cold Spring Harbor Press, Cold Spring Harbor, NY (1980), which is incorporated herein by reference in its entirety); the trp, recA, heat shock, and lacZ promoters of *E. coli* and the SV40 early promoter (Benoist *et al. Nature*, 1981, 290, 304-310, which is incorporated herein by reference in its entirety). Additional promoters include, but are not limited to, mouse mammary tumor virus, long terminal repeat of human immunodeficiency virus, maloney virus, cytomegalovirus immediate early promoter, Epstein Barr virus, Rous sarcoma

virus, human actin, human myosin, human hemoglobin, human muscle creatine, and human metallothionein.

Additional regulatory sequences can also be included in preferred vectors. Preferred examples of suitable regulatory sequences are represented by the Shine-Dalgarno of the replicase gene of the phage MS-2 and of the gene cII of bacteriophage lambda. The Shine-Dalgarno sequence may be directly followed by DNA encoding Con-218 and result in the expression of the mature Con-218 protein.

Moreover, suitable expression vectors can include an appropriate marker that allows the screening of the transformed host cells. The transformation of the selected host is carried out using any one of the various techniques well known to the expert in the art and described in Sambrook *et al.*, *supra*.

An origin of replication can also be provided either by construction of the vector to include an exogenous origin or may be provided by the host cell chromosomal replication mechanism. If the vector is integrated into the host cell chromosome, the latter may be sufficient. Alternatively, rather than using vectors which contain viral origins of replication, one skilled in the art can transform mammalian cells by the method of co-transformation with a selectable marker and Con-218 DNA. An example of a suitable marker is dihydrofolate reductase (DHFR) or thymidine kinase (*see*, U.S. Patent No. 4,399,216).

Nucleotide sequences encoding Con-218 may be recombined with vector DNA in accordance with conventional techniques, including blunt-ended or staggered-ended termini for ligation, restriction enzyme digestion to provide appropriate termini, filling in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and ligation with appropriate ligases. Techniques for such manipulation are disclosed by Sambrook *et al.*, *supra* and are well known in the art. Methods for construction of mammalian expression vectors are disclosed in, for example, Okayama *et al.*, *Mol. Cell. Biol.*, **1983**, 3, 280, Cosman *et al.*, *Mol. Immunol.*, **1986**, 23, 935, Cosman *et al.*, *Nature*, **1984**, 312, 768, EP-A-0367566, and WO 91/18982, each of which is incorporated herein by reference in its entirety.

**Host cells**

According to another aspect of the invention, host cells are provided, including prokaryotic and eukaryotic cells, comprising a polynucleotide of the invention (or vector of the invention) in a manner that permits expression of the encoded Con-218 polypeptide.

Polynucleotides of the invention may be introduced into the host cell as part of a circular plasmid, or as linear DNA comprising an isolated protein coding region or a viral vector.

Methods for introducing DNA into the host cell that are well known and routinely practiced in the art include transformation, transfection, electroporation, nuclear injection, or fusion with carriers such as liposomes, micelles, ghost cells, and protoplasts. Expression systems of the invention include bacterial, yeast, fungal, plant, insect, invertebrate, vertebrate, and mammalian cells systems.

The invention provides host cells that are transformed or transfected (stably or transiently) with polynucleotides of the invention or vectors of the invention. As stated above, such host cells are useful for amplifying the polynucleotides and also for expressing the Con-218 polypeptide or fragment thereof encoded by the polynucleotide.

In still another related embodiment, the invention provides a method for producing a Con-218 polypeptide (or fragment thereof) comprising the steps of growing a host cell of the invention in a nutrient medium and isolating the polypeptide or variant thereof from the cell or the medium. Because Con-218 is a seven transmembrane receptor, it will be appreciated that, for some applications, such as certain activity assays, the preferable isolation may involve isolation of cell membranes containing the polypeptide embedded therein, whereas for other applications a more complete isolation may be preferable.

According to some aspects of the present invention, transformed host cells having an expression vector comprising any of the nucleic acid molecules described above are provided. Expression of the nucleotide sequence occurs when the expression vector is introduced into an appropriate host cell. Suitable host cells for expression of the polypeptides of the invention include, but are not limited to, prokaryotes, yeast, and eukaryotes. If a prokaryotic expression



vector is employed, then the appropriate host cell would be any prokaryotic cell capable of expressing the cloned sequences. Suitable prokaryotic cells include, but are not limited to, bacteria of the genera *Escherichia*, *Bacillus*, *Salmonella*, *Pseudomonas*, *Streptomyces*, and *Staphylococcus*.

If an eukaryotic expression vector is employed, then the appropriate host cell would be any eukaryotic cell capable of expressing the cloned sequence. Preferably, eukaryotic cells are cells of higher eukaryotes. Suitable eukaryotic cells include, but are not limited to, non-human mammalian tissue culture cells and human tissue culture cells. Preferred host cells include, but are not limited to, insect cells, HeLa cells, Chinese hamster ovary cells (CHO cells), African green monkey kidney cells (COS cells), human 293 cells, and murine 3T3 fibroblasts. Propagation of such cells in cell culture has become a routine procedure (*see*; Tissue Culture, Academic Press, Kruse and Patterson, eds. (1973), which is incorporated herein by reference in its entirety).

In addition, a yeast host may be employed as a host cell. Preferred yeast cells include, but are not limited to, the genera *Saccharomyces*, *Pichia*, and *Kluveromyces*. Preferred yeast hosts are *S. cerevisiae* and *P. pastoris*. Preferred yeast vectors can contain an origin of replication sequence from a 2T yeast plasmid, an autonomously replication sequence (ARS), a promoter region, sequences for polyadenylation, sequences for transcription termination, and a selectable marker gene. Shuttle vectors for replication in both yeast and *E. coli* are also included herein.

Alternatively, insect cells may be used as host cells. In a preferred embodiment, the polypeptides of the invention are expressed using a baculovirus expression system (*see*, Luckow *et al.*, *Bio/Technology*, **1988**, 6, 47, Baculovirus Expression Vectors: A Laboratory Manual, O'Rielly *et al.* (Eds.), W.H. Freeman and Company, New York, **1992**, and U.S. Patent No. 4,879,236, each of which is incorporated herein by reference in its entirety). In addition, the MAXBAC™ complete baculovirus expression system (Invitrogen) can, for example, be used for production in insect cells.

Host cells of the invention are a valuable source of immunogen for development of antibodies specifically immunoreactive with Con-218. Host cells of the invention are also useful in methods for the large-scale production of Con-218 polypeptides wherein the cells are grown in a suitable culture medium and the desired polypeptide products are isolated from the cells, or from the medium in which the cells are grown, by purification methods known in the art, *e.g.*, conventional chromatographic methods including immunoaffinity chromatography, receptor affinity chromatography, hydrophobic interaction chromatography, lectin affinity chromatography, size exclusion filtration, cation or anion exchange chromatography, high pressure liquid chromatography (HPLC), reverse phase HPLC, and the like. Still other methods of purification include those methods wherein the desired protein is expressed and purified as a fusion protein having a specific tag, label, or chelating moiety that is recognized by a specific binding partner or agent. The purified protein can be cleaved to yield the desired protein, or can be left as an intact fusion protein. Cleavage of the fusion component may produce a form of the desired protein having additional amino acid residues as a result of the cleavage process.

Knowledge of Con-218 DNA sequences allows for modification of cells to permit, or increase, expression of endogenous Con-218. Cells can be modified (*e.g.*, by homologous recombination) to provide increased expression by replacing, in whole or in part, the naturally occurring Con-218 promoter with all or part of a heterologous promoter so that the cells express Con-218 at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to endogenous Con-218 encoding sequences. (See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955.) It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (*e.g.*, *ada*, *dhfr*, and the multifunctional CAD gene which encodes carbamoyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the Con-218 coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the Con-218 coding sequences in the cells.

**Knock-outs**

The DNA sequence information provided by the present invention also makes possible the development (*e.g.*, by homologous recombination or "knock-out" strategies; see Capecchi, *Science* 244:1288-1292 (1989), which is incorporated herein by reference) of animals that fail to express functional Con-218 or that express a variant of Con-218. Such animals (especially small laboratory animals such as rats, rabbits, and mice) are useful as models for studying the *in vivo* activities of Con-218 and modulators of Con-218.

**Antisense**

Also made available by the invention are anti-sense polynucleotides that recognize and hybridize to polynucleotides encoding Con-218. Full-length and fragment anti-sense polynucleotides are provided. Fragment antisense molecules of the invention include (i) those that specifically recognize and hybridize to Con-218 RNA (as determined by sequence comparison of DNA encoding Con-218 to DNA encoding other known molecules). Identification of sequences unique to Con-218 encoding polynucleotides can be deduced through use of any publicly available sequence database, and/or through use of commercially available sequence comparison programs. After identification of the desired sequences, isolation through restriction digestion or amplification using any of the various polymerase chain reaction techniques well known in the art can be performed. Anti-sense polynucleotides are particularly relevant to regulating expression of Con-218 by those cells expressing Con-218 mRNA.

Antisense nucleic acids (preferably 10 to 30 base-pair oligonucleotides) capable of specifically binding to Con-218 expression control sequences or Con-218 RNA are introduced into cells (*e.g.*, by a viral vector or colloidal dispersion system such as a liposome). The antisense nucleic acid binds to the Con-218 target nucleotide sequence in the cell and prevents transcription and/or translation of the target sequence. Phosphorothioate and methylphosphonate antisense oligonucleotides are specifically contemplated for therapeutic use by the invention. The antisense oligonucleotides may be further modified by adding poly-L-lysine, transferrin polylysine, or cholesterol moieties at their 5' end. Suppression of Con-218 expression at either

the transcriptional or translational level is useful to generate cellular or animal models for diseases/conditions characterized by aberrant Con-218 expression.

Antisense oligonucleotides, or fragments of the sequence of SEQ ID NO: 1 or 3, or sequences complementary or homologous thereto, derived from the nucleotide sequences of the present invention encoding Con-218 are useful as diagnostic tools for probing gene expression in various tissues. For example, tissue can be probed *in situ* with oligonucleotide probes carrying detectable groups by conventional autoradiography techniques to investigate native expression of this enzyme or pathological conditions relating thereto. Antisense oligonucleotides are preferably directed to regulatory regions of a sequence of SEQ ID NO: 1 or 3, or mRNA corresponding thereto, including, but not limited to, the initiation codon, TATA box, enhancer sequences, and the like.

### **Transcription factors**

The Con-218 sequences taught in the present invention facilitate the design of novel transcription factors for modulating Con-218 expression in native cells and animals, and cells transformed or transfected with Con-218 polynucleotides. For example, the Cys<sub>2</sub>-His<sub>2</sub> zinc finger proteins, which bind DNA via their zinc finger domains, have been shown to be amenable to structural changes that lead to the recognition of different target sequences. These artificial zinc finger proteins recognize specific target sites with high affinity and low dissociation constants, and are able to act as gene switches to modulate gene expression. Knowledge of the particular Con-218 target sequence of the present invention facilitates the engineering of zinc finger proteins specific for the target sequence using known methods such as a combination of structure-based modeling and screening of phage display libraries (Segal *et al.*, Proc. Natl. Acad. Sci. (USA) 96:2758-2763 (1999); Liu *et al.*, Proc. Natl. Acad. Sci. (USA) 94:5525-5530 (1997); Greisman *et al.*, Science 275:657-661 (1997); Choo *et al.*, J. Mol. Biol. 273:525-532 (1997)). Each zinc finger domain usually recognizes three or more base pairs. Since a recognition sequence of 18 base pairs is generally sufficient in length to render it unique in any known genome, a zinc finger protein consisting of 6 tandem repeats of zinc fingers would be expected to

ensure specificity for a particular sequence (Segal *et al.*) The artificial zinc finger repeats, designed based on Con-218 sequences, are fused to activation or repression domains to promote or suppress Con-218 expression (Liu *et al.*) Alternatively, the zinc finger domains can be fused to the TATA box-binding factor (TBP) with varying lengths of linker region between the zinc finger peptide and the TBP to create either transcriptional activators or repressors (Kim *et al.*, Proc. Natl. Acad. Sci. (USA) 94:3616-3620 (1997). Such proteins and polynucleotides that encode them, have utility for modulating Con-218 expression *in vivo* in both native cells, animals and humans; and/or cells transfected with Con-218-encoding sequences. The novel transcription factor can be delivered to the target cells by transfecting constructs that express the transcription factor (gene therapy), or by introducing the protein. Engineered zinc finger proteins can also be designed to bind RNA sequences for use in therapeutics as alternatives to antisense or catalytic RNA methods (McColl *et al.*, Proc. Natl. Acad. Sci. (USA) 96:9521-9526 (1997); Wu *et al.*, Proc. Natl. Acad. Sci. (USA) 92:344-348 (1995)). The present invention contemplates methods of designing such transcription factors based on the gene sequence of the invention, as well as customized zinc finger proteins, that are useful to modulate Con-218 expression in cells (native or transformed) whose genetic complement includes these sequences.

### Polypeptides

The invention also provides purified and isolated mammalian Con-218 polypeptides encoded by a polynucleotide of the invention. Presently preferred is a human Con-218 polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO: 2 and SEQ ID NO:4, or fragments thereof comprising an epitope specific to the polypeptide. By "epitope specific to" is meant a portion of the nGPCR receptor that is recognizable by an antibody that is specific for the nGPCR, as defined in detail below.

Although the sequences provided are particular human sequences, the invention is intended to include within its scope other human allelic variants; non-human mammalian forms of Con-218, and other vertebrate forms of Con-218.

It will be appreciated that extracellular epitopes are particularly useful for generating and screening for antibodies and other binding compounds that bind to receptors such as Con-218. Thus, in another preferred embodiment, the invention provides a purified and isolated polypeptide comprising at least one extracellular domain (e.g., the N-terminal extracellular domain or one of the three extracellular loops) of Con-218. Purified and isolated polypeptides comprising the N-terminal extracellular domain of Con-218 are highly preferred. Also preferred is a purified and isolated polypeptide comprising a Con-218 fragment selected from the group consisting of the N-terminal extracellular domain of Con-218, transmembrane domains of Con-218, an extracellular loop connecting transmembrane domains of Con-218, an intracellular loop connecting transmembrane domains of Con-218, the C-terminal cytoplasmic region of Con-218, and fusions thereof. Such fragments may be continuous portions of the native receptor. However, it will also be appreciated that knowledge of the Con-218 gene and protein sequences as provided herein permits recombining of various domains that are not contiguous in the native protein. Using a FORTRAN computer program called "tmrest.all" [Parodi *et al.*, Comput. Appl. Biosci. 5:527-535 (1994)], Con-218 was shown to contain transmembrane-spanning domains.

The invention also embraces polypeptides that have at least 99%, at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 65%, at least 60%, at least 55% or at least 50% identity and/or homology to the preferred polypeptide of the invention. Percent amino acid sequence "identity" with respect to the preferred polypeptide of the invention is defined herein as the percentage of amino acid residues in the candidate sequence that are identical with the residues in the Con-218 sequence after aligning both sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Percent sequence "homology" with respect to the preferred polypeptide of the invention is defined herein as the percentage of amino acid residues in the candidate sequence that are identical with the residues in the Con-218 sequence after aligning the sequences and introducing gaps, if necessary, to achieve

the maximum percent sequence identity, and also considering any conservative substitutions as part of the sequence identity.

In one aspect, percent homology is calculated as the percentage of amino acid residues in the smaller of two sequences which align with identical amino acid residue in the sequence being compared, when four gaps in a length of 100 amino acids may be introduced to maximize alignment [Dayhoff, in Atlas of Protein Sequence and Structure, Vol. 5, p. 124, National Biochemical Research Foundation, Washington, D.C. (1972), incorporated herein by reference].

Polypeptides of the invention may be isolated from natural cell sources or may be chemically synthesized, but are preferably produced by recombinant procedures involving host cells of the invention. Use of mammalian host cells is expected to provide for such post-translational modifications (*e.g.*, glycosylation, truncation, lipidation, and phosphorylation) as may be needed to confer optimal biological activity on recombinant expression products of the invention. Glycosylated and non-glycosylated forms of Con-218 polypeptides are embraced by the invention.

The invention also embraces variant (or analog) Con-218 polypeptides. In one example, insertion variants are provided wherein one or more amino acid residues supplement a Con-218 amino acid sequence. Insertions may be located at either or both termini of the protein, or may be positioned within internal regions of the Con-218 amino acid sequence. Insertional variants with additional residues at either or both termini can include, for example, fusion proteins and proteins including amino acid tags or labels.

Insertion variants include Con-218 polypeptides wherein one or more amino acid residues are added to a Con-218 acid sequence or to a biologically active fragment thereof.

Variant products of the invention also include mature Con-218 products, *i.e.*, Con-218 products wherein leader or signal sequences are removed, with additional amino terminal residues. The additional amino terminal residues may be derived from another protein, or may include one or more residues that are not identifiable as being derived from specific proteins. Con-218 products with an additional methionine residue at position -1 (Met<sup>-1</sup>-Con-218) are

contemplated, as are variants with additional methionine and lysine residues at positions -2 and -1 (Met<sup>-2</sup>-Lys<sup>-1</sup>-Con-218). Variants of Con-218 with additional Met, Met-Lys, Lys residues (or one or more basic residues in general) are particularly useful for enhanced recombinant protein production in bacterial host cells.

The invention also embraces Con-218 variants having additional amino acid residues that result from use of specific expression systems. For example, use of commercially available vectors that express a desired polypeptide as part of a glutathione-S-transferase (GST) fusion product provides the desired polypeptide having an additional glycine residue at position -1 after cleavage of the GST component from the desired polypeptide. Variants that result from expression in other vector systems are also contemplated.

Insertional variants also include fusion proteins wherein the amino terminus and/or the carboxy terminus of Con-218 is/are fused to another polypeptide.

In another aspect, the invention provides deletion variants wherein one or more amino acid residues in a Con-218 polypeptide are removed. Deletions can be effected at one or both termini of the Con-218 polypeptide, or with removal of one or more non-terminal amino acid residues of Con-218. Deletion variants, therefore, include all fragments of a Con-218 polypeptide.

The invention also embraces polypeptide fragments of sequences selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:4, wherein the fragments maintain biological (*e.g.*, ligand binding and/or intracellular signaling) immunological properties of a Con-218 polypeptide.

In one preferred embodiment of the invention, an isolated nucleic acid molecule comprises a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence homologous to a sequence selected from the group consisting of: SEQ ID NO:2 and SEQ ID NO:4, and fragments thereof, wherein the nucleic acid molecule encoding at least a portion of Con-218. In a more preferred embodiment, the isolated nucleic acid molecule comprises a



sequence that encodes a polypeptide comprising a sequence selected from the group consisting of: SEQ ID NO:2 and SEQ ID NO:4, and fragments thereof.

As used in the present invention, polypeptide fragments comprise at least 5, 10, 15, 20, 25, 30, 35, or 40 consecutive amino acids of SEQ ID NO:2 or SEQ ID NO:4. Preferred polypeptide fragments display antigenic properties unique to, or specific for, human Con-218 and its allelic and species homologs. Fragments of the invention having the desired biological and immunological properties can be prepared by any of the methods well known and routinely practiced in the art.

In still another aspect, the invention provides substitution variants of Con-218 polypeptides. Substitution variants include those polypeptides wherein one or more amino acid residues of a Con-218 polypeptide are removed and replaced with alternative residues. In one aspect, the substitutions are conservative in nature; however, the invention embraces substitutions that are also non-conservative. Conservative substitutions for this purpose may be defined as set out in Tables 1, 2, or 3, below.

Variant polypeptides include those wherein conservative substitutions have been introduced by modification of polynucleotides encoding polypeptides of the invention. Amino acids can be classified according to physical properties and contribution to secondary and tertiary protein structure. A conservative substitution is recognized in the art as a substitution of one amino acid for another amino acid that has similar properties. Exemplary conservative substitutions are set out in Table 1 (from WO 97/09433, page 10, published March 13, 1997 (PCT/GB96/02197, filed 9/6/96), immediately below.

<b>Table 1</b>	
<b>Conservative Substitutions I</b>	
<b><u>SIDE CHAIN</u></b>	<b><u>AMINO ACID</u></b>
<b><u>CHARACTERISTIC</u></b>	
Aliphatic	
Non-polar	G A P I L V

Polar - uncharged	C S T M
	N Q
Polar - charged	D E
	K R
Aromatic	H F W Y
Other	N Q D E

Alternatively, conservative amino acids can be grouped as described in Lehninger, [Biochemistry, Second Edition; Worth Publishers, Inc. NY, NY (1975), pp.71-77] as set out in Table 2, below.

**Table 2**  
**Conservative Substitutions II**

<u>SIDE CHAIN</u> <u>CHARACTERISTIC</u>	<u>AMINO ACID</u>
Non-polar (hydrophobic)	
A. Aliphatic:	A L I V
P	
B. Aromatic:	F W
C. Sulfur-containing:	M
D. Borderline:	G
Uncharged-polar	
A. Hydroxyl:	S T Y
B. Amides:	N Q
C. Sulfhydryl:	C
D. Borderline:	G
Positively Charged (Basic):	K R H
Negatively Charged (Acidic):	D E

As still another alternative, exemplary conservative substitutions are set out in Table 3, below.

**Table 3**  
**Conservative Substitutions III**

<b>Original</b>	<b>Exemplary Substitution</b>
-----------------	-------------------------------

**Residue**

Ala (A)	Val, Leu, Ile
Arg (R)	Lys, Gln, Asn
Asn (N)	Gln, His, Lys, Arg
Asp (D)	Glu
Cys (C)	Ser
Gln (Q)	Asn
Glu (E)	Asp
His (H)	Asn, Gln, Lys, Arg
Ile (I)	Leu, Val, Met, Ala, Phe,
Leu (L)	Ile, Val, Met, Ala, Phe
Lys (K)	Arg, Gln, Asn
Met (M)	Leu, Phe, Ile
Phe (F)	Leu, Val, Ile, Ala
Pro (P)	Gly
Ser (S)	Thr
Thr (T)	Ser
Trp (W)	Tyr
Tyr (Y)	Trp, Phe, Thr, Ser
Val (V)	Ile, Leu, Met, Phe, Ala

It should be understood that the definition of polypeptides of the invention is intended to include polypeptides bearing modifications other than insertion, deletion, or substitution of amino acid residues. By way of example, the modifications may be covalent in nature, and include for example, chemical bonding with polymers, lipids, other organic, and inorganic moieties. Such derivatives may be prepared to increase circulating half-life of a polypeptide, or may be designed to improve the targeting capacity of the polypeptide for desired cells, tissues, or organs. Similarly, the invention further embraces Con-218 polypeptides that have been covalently modified to include one or more water-soluble polymer attachments such as polyethylene glycol, polyoxyethylene glycol, or polypropylene glycol. Variants that display ligand binding properties of native Con-218 and are expressed at higher levels, as well as

variants that provide for constitutively active receptors, are particularly useful in assays of the invention; the variants are also useful in providing cellular, tissue and animal models of diseases/conditions characterized by aberrant Con-218 activity.

In a related embodiment, the present invention provides compositions comprising purified polypeptides of the invention. Preferred compositions comprise, in addition to the polypeptide of the invention, a pharmaceutically acceptable (*i.e.*, sterile and non-toxic) liquid, semisolid, or solid diluent that serves as a pharmaceutical vehicle, excipient, or medium. Any diluent known in the art may be used. Exemplary diluents include, but are not limited to, water, saline solutions, polyoxyethylene sorbitan monolaurate, magnesium stearate, methyl- and propylhydroxybenzoate, talc, alginates, starches, lactose, sucrose, dextrose, sorbitol, mannitol, glycerol, calcium phosphate, mineral oil, and cocoa butter.

Variants that display ligand binding properties of native Con-218 and are expressed at higher levels, as well as variants that provide for constitutively active receptors, are particularly useful in assays of the invention; the variants are also useful in assays of the invention and in providing cellular, tissue and animal models of diseases/conditions characterized by aberrant Con-218 activity.

The G protein-coupled receptor functions through a specific heterotrimeric guanine-nucleotide-binding regulatory protein (G-protein) coupled to the intracellular portion of the G protein-coupled receptor molecule. Accordingly, the G protein-coupled receptor has a specific affinity to G protein. G proteins specifically bind to guanine nucleotides. Isolation of G proteins provides a means to isolate guanine nucleotides. G Proteins may be isolated using commercially available anti-G protein antibodies or isolated G protein-coupled receptors. Similarly, G proteins may be detected in a sample isolated using commercially available detectable anti-G protein antibodies or isolated G protein-coupled receptors.

According to the present invention, the isolated Con-218 proteins of the present invention are useful to isolate and purify G proteins from samples such as cell lysates. Example 15 below sets forth an example of isolation of G proteins using isolated Con-218 proteins. Such

methodology may be used in place of the use of commercially available anti-G protein antibodies which are used to isolate G proteins. Moreover, G proteins may be detected using Con-218 proteins in place of commercially available detectable anti-G protein antibodies. Since Con-218 proteins specifically bind to G proteins, they can be employed in any specific use where G protein specific affinity is required such as those uses where commercially available anti-G protein antibodies are employed.

### Antibodies

Also comprehended by the present invention are antibodies (*e.g.*, monoclonal and polyclonal antibodies, single chain antibodies, chimeric antibodies, bifunctional/bispecific antibodies, humanized antibodies, human antibodies, and complementary determining region (CDR)-grafted antibodies, including compounds which include CDR sequences which specifically recognize a polypeptide of the invention) specific for Con-218 or fragments thereof. Preferred antibodies of the invention are human antibodies that are produced and identified according to methods described in WO93/11236, published June 20, 1993, which is incorporated herein by reference in its entirety. Antibody fragments, including Fab, Fab', F(ab')<sub>2</sub>, and F<sub>v</sub>, are also provided by the invention. The term "specific for," when used to describe antibodies of the invention, indicates that the variable regions of the antibodies of the invention recognize and bind Con-218 polypeptides exclusively (*i.e.*, are able to distinguish Con-218 polypeptides from other known GPCR polypeptides by virtue of measurable differences in binding affinity, despite the possible existence of localized sequence identity, homology, or similarity between Con-218 and such polypeptides). It will be understood that specific antibodies may also interact with other proteins (for example, *S. aureus* protein A or other antibodies in ELISA techniques) through interactions with sequences outside the variable region of the antibodies, and, in particular, in the constant region of the molecule. Screening assays to determine binding specificity of an antibody of the invention are well known and routinely practiced in the art. For a comprehensive discussion of such assays, see Harlow *et al.* (Eds.), Antibodies A Laboratory Manual; Cold Spring Harbor Laboratory; Cold Spring Harbor, NY (1988), Chapter 6.

Antibodies that recognize and bind fragments of the Con-218 polypeptides of the invention are also contemplated, provided that the antibodies are specific for Con-218 polypeptides.

Antibodies of the invention can be produced using any method well known and routinely practiced in the art.

The invention provides an antibody that is specific for the Con-218 of the invention. Antibody specificity is described in greater detail below. However, it should be emphasized that antibodies that can be generated from polypeptides that have previously been described in the literature and that are capable of fortuitously cross-reacting with Con-218 (*e.g.*, due to the fortuitous existence of a similar epitope in both polypeptides) are considered “cross-reactive” antibodies. Such cross-reactive antibodies are not antibodies that are “specific” for Con-218. The determination of whether an antibody is specific for Con-218 or is cross-reactive with another known receptor is made using any of several assays, such as Western blotting assays, that are well known in the art. For identifying cells that express Con-218 and also for modulating Con-218-ligand binding activity, antibodies that specifically bind to an extracellular epitope of the Con-218 are preferred.

In one preferred variation, the invention provides monoclonal antibodies. Hybridomas that produce such antibodies also are intended as aspects of the invention. In yet another variation, the invention provides a humanized antibody. Humanized antibodies are useful for *in vivo* therapeutic indications.

In another variation, the invention provides a cell-free composition comprising polyclonal antibodies, wherein at least one of the antibodies is an antibody of the invention specific for Con-218. Antisera isolated from an animal is an exemplary composition, as is a composition comprising an antibody fraction of an antisera that has been resuspended in water or in another diluent, excipient, or carrier.

In still another related embodiment, the invention provides an anti-idiotypic antibody specific for an antibody that is specific for Con-218.

It is well known that antibodies contain relatively small antigen binding domains that can be isolated chemically or by recombinant techniques. Such domains are useful Con-218 binding molecules themselves, and also may be reintroduced into human antibodies, or fused to toxins or other polypeptides. Thus, in still another embodiment, the invention provides a polypeptide comprising a fragment of a Con-218-specific antibody, wherein the fragment and the polypeptide bind to the Con-218. By way of non-limiting example, the invention provides polypeptides that are single chain antibodies and CDR-grafted antibodies.

Non-human antibodies may be humanized by any of the methods known in the art. In one method, the non-human CDRs are inserted into a human antibody or consensus antibody framework sequence. Further changes can then be introduced into the antibody framework to modulate affinity or immunogenicity.

Antibodies of the invention are useful for, *e.g.*, therapeutic purposes (by modulating activity of Con-218), diagnostic purposes to detect or quantitate Con-218, and purification of Con-218. Kits comprising an antibody of the invention for any of the purposes described herein are also comprehended. In general, a kit of the invention also includes a control antigen for which the antibody is immunospecific.

### Compositions

Mutations in the Con-218 gene that result in loss of normal function of the Con-218 gene product underlie Con-218-related human disease states. The invention comprehends gene therapy to restore Con-218 activity to treat those disease states. Delivery of a functional Con-218 gene to appropriate cells is effected *ex vivo*, *in situ*, or *in vivo* by use of vectors, and more particularly viral vectors (*e.g.*, adenovirus, adeno-associated virus, or a retrovirus), or *ex vivo* by use of physical DNA transfer methods (*e.g.*, liposomes or chemical treatments). See, for example, Anderson, *Nature*, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, *Science*, 244: 1275-1281 (1989); Verma, *Scientific American*: 68-84 (1990); and Miller, *Nature*, 357: 455-460 (1992). Alternatively, it is contemplated that in other human disease states, preventing the expression of, or inhibiting the

activity of, Con-218 will be useful in treating disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of Con-218.

Another aspect of the present invention is directed to compositions, including pharmaceutical compositions, comprising any of the nucleic acid molecules or recombinant expression vectors described above and an acceptable carrier or diluent. Preferably, the carrier or diluent is pharmaceutically acceptable. Suitable carriers are described in the most recent edition of *Remington's Pharmaceutical Sciences*, A. Osol, a standard reference text in this field, which is incorporated herein by reference in its entirety. Preferred examples of such carriers or diluents include, but are not limited to, water, saline, Ringer's solution, dextrose solution, and 5% human serum albumin. Liposomes and nonaqueous vehicles such as fixed oils may also be used. The formulations are sterilized by commonly used techniques.

Also within the scope of the invention are compositions comprising polypeptides, polynucleotides, or antibodies of the invention that have been formulated with, *e.g.*, a pharmaceutically acceptable carrier.

The invention also provides methods of using antibodies of the invention. For example, the invention provides a method for modulating ligand binding of a Con-218 comprising the step of contacting the Con-218 with an antibody specific for the Con-218, under conditions wherein the antibody binds the receptor.

GPCRs that may be expressed in the brain, such as Con-218, provide an indication that aberrant Con-218 signaling activity may correlate with one or more neurological or psychological disorders. The invention also provides a method for treating a neurological or psychiatric disorder comprising the step of administering to a mammal in need of such treatment an amount of an antibody-like polypeptide of the invention that is sufficient to modulate ligand binding to a Con-218 in neurons of the mammal. Con-218 transcripts were found in the central nervous system, with highest levels in all lobes of the cerebral cortex, cerebellum, corpus callosum, amygdala, substantia nigra, and the nucleus caudalis. Con-218 transcripts were also detected in the lymph node, testis, and pituitary.



**Kits**

The present invention is also directed to kits, including pharmaceutical kits. The kits can comprise any of the nucleic acid molecules described above, any of the polypeptides described above, or any antibody which binds to a polypeptide of the invention as described above, as well as a negative control. The kit preferably comprises additional components, such as, for example, instructions, solid support, reagents helpful for quantification, and the like.

In another aspect, the invention features methods for detection of a polypeptide in a sample as a diagnostic tool for diseases or disorders, wherein the method comprises the steps of: (a) contacting the sample with a nucleic acid probe which hybridizes under hybridization assay conditions to a nucleic acid target region of a polypeptide having the sequence of a sequence selected from the group consisting of: SEQ ID NO:2 and SEQ ID NO:4, said probe comprising the nucleic acid sequence encoding the polypeptide, fragments thereof, and the complements of the sequences and fragments; and (b) detecting the presence or amount of the probe:target region hybrid as an indication of the disease.

In preferred embodiments of the invention, the disease is selected from the group consisting of thyroid disorders (*e.g.* thyreotoxicosis, myxoedema); renal failure; inflammatory conditions (*e.g.*, Crohn's disease); diseases related to cell differentiation and homeostasis; rheumatoid arthritis; autoimmune disorders; movement disorders; CNS disorders (*e.g.*, pain including migraine; stroke; psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, anxiety, generalized anxiety disorder, post-traumatic-stress disorder, depression, bipolar disorder, delirium, dementia, severe mental retardation; dyskinesias, such as Huntington's disease or Tourette's Syndrome; attention disorders including ADD and ADHD, and degenerative disorders such as Parkinson's, Alzheimer's; movement disorders, including ataxias, supranuclear palsy, *etc.*); infections, such as viral infections caused by HIV-1 or HIV-2; metabolic and cardiovascular diseases and disorders (*e.g.*, type 2 diabetes, obesity, anorexia, hypotension, hypertension, thrombosis, myocardial infarction, cardiomyopathies, atherosclerosis, *etc.*); proliferative diseases and cancers (*e.g.*, different cancers such as breast,

colon, lung, *etc.*, and hyperproliferative disorders such as psoriasis, prostate hyperplasia, *etc.*); hormonal disorders (*e.g.*, male/female hormonal replacement, polycystic ovarian syndrome, alopecia, *etc.*); and sexual dysfunction, among others.

As described above and in Example 4 below, the gene encoding Con-218 has been detected in brain tissue indicating that these Con-218 proteins are neuroreceptors. Kits may be designed to detect either expression of polynucleotides encoding these proteins or the proteins themselves in order to identify tissue as being neurological. For example, oligonucleotide hybridization kits can be provided which include a container having an oligonucleotide probe specific for the Con-218-specific DNA and optionally, containers with positive and negative controls and/or instructions. Similarly, PCR kits can be provided which include a container having primers specific for the Con-218-specific sequences, DNA and optionally, containers with size markers, positive and negative controls and/or instructions.

Hybridization conditions should be such that hybridization occurs only with the genes in the presence of other nucleic acid molecules. Under stringent hybridization conditions only highly complementary nucleic acid sequences hybridize. Preferably, such conditions prevent hybridization of nucleic acids having 1 or 2 mismatches out of 20 contiguous nucleotides. Such conditions are defined *supra*.

The diseases for which detection of genes in a sample could be diagnostic include diseases in which nucleic acid (DNA and/or RNA) is amplified in comparison to normal cells. By "amplification" is meant increased numbers of DNA or RNA in a cell compared with normal cells.

The diseases that could be diagnosed by detection of nucleic acid in a sample preferably include central nervous system and metabolic diseases. The test samples suitable for nucleic acid probing methods of the present invention include, for example, cells or nucleic acid extracts of cells, or biological fluids. The samples used in the above-described methods will vary based on the assay format, the detection method and the nature of the tissues, cells or extracts to be

assayed. Methods for preparing nucleic acid extracts of cells are well known in the art and can be readily adapted in order to obtain a sample that is compatible with the method utilized.

Alternatively, immunoassay kits can be provided which have containers container having antibodies specific for the Con-218 protein and optionally, containers with positive and negative controls and/or instructions.

Kits may also be provided useful in the identification of GPCR binding partners such as natural ligands or modulators (agonists or antagonists). Substances useful for treatment of disorders or diseases preferably show positive results in one or more *in vitro* assays for an activity corresponding to treatment of the disease or disorder in question. Substances that modulate the activity of the polypeptides preferably include, but are not limited to, antisense oligonucleotides, agonists and antagonists, and inhibitors of protein kinases.

#### **Methods of inducing immune response**

Another aspect of the present invention is directed to methods of inducing an immune response in a mammal against a polypeptide of the invention by administering to the mammal an amount of the polypeptide sufficient to induce an immune response. The amount will be dependent on the animal species, size of the animal, and the like but can be determined by those skilled in the art.

#### **Methods of identifying ligands**

The invention also provides assays to identify compounds that bind Con-218. One such assay comprises the steps of: (a) contacting a composition comprising a Con-218 with a compound suspected of binding Con-218; and (b) measuring binding between the compound and Con-218. In one variation, the composition comprises a cell expressing Con-218 on its surface. In another variation, isolated Con-218 or cell membranes comprising Con-218 are employed. The binding may be measured directly, *e.g.*, by using a labeled compound, or may be measured indirectly by several techniques, including measuring intracellular signaling of Con-218 induced by the compound (or measuring changes in the level of Con-218 signaling).

Specific binding molecules, including natural ligands and synthetic compounds, can be identified or developed using isolated or recombinant Con-218 products, Con-218 variants, or preferably, cells expressing such products. Binding partners are useful for purifying Con-218 products and detection or quantification of Con-218 products in fluid and tissue samples using known immunological procedures. Binding molecules are also manifestly useful in modulating (*i.e.*, blocking, inhibiting or stimulating) biological activities of Con-218, especially those activities involved in signal transduction.

The DNA and amino acid sequence information provided by the present invention also makes possible identification of binding partner compounds with which a Con-218 polypeptide or polynucleotide will interact. Methods to identify binding partner compounds include solution assays, *in vitro* assays wherein Con-218 polypeptides are immobilized, and cell-based assays. Identification of binding partner compounds of Con-218 polypeptides provides candidates for therapeutic or prophylactic intervention in pathologies associated with Con-218 normal and aberrant biological activity.

The invention includes several assay systems for identifying Con-218 binding partners. In solution assays, methods of the invention comprise the steps of (a) contacting a Con-218 polypeptide with one or more candidate binding partner compounds and (b) identifying the compounds that bind to the Con-218 polypeptide. Identification of the compounds that bind the Con-218 polypeptide can be achieved by isolating the Con-218 polypeptide/binding partner complex, and separating the binding partner compound from the Con-218 polypeptide. An additional step of characterizing the physical, biological, and/or biochemical properties of the binding partner compound is also comprehended in another embodiment of the invention. In one aspect, the Con-218 polypeptide/binding partner complex is isolated using an antibody immunospecific for either the Con-218 polypeptide or the candidate binding partner compound.

In still other embodiments, either the Con-218 polypeptide or the candidate binding partner compound comprises a label or tag that facilitates its isolation, and methods of the invention to identify binding partner compounds include a step of isolating the Con-218

polypeptide/binding partner complex through interaction with the label or tag. An exemplary tag of this type is a poly-histidine sequence, generally around six histidine residues, that permits isolation of a compound so labeled using nickel chelation. Other labels and tags, such as the FLAG® tag (Eastman Kodak, Rochester, NY), well known and routinely used in the art, are embraced by the invention.

In one variation of an *in vitro* assay, the invention provides a method comprising the steps of (a) contacting an immobilized Con-218 polypeptide with a candidate binding partner compound and (b) detecting binding of the candidate compound to the Con-218 polypeptide. In an alternative embodiment, the candidate binding partner compound is immobilized and binding of Con-218 is detected. Immobilization is accomplished using any of the methods well known in the art, including covalent bonding to a support, a bead, or a chromatographic resin, as well as non-covalent, high affinity interactions such as antibody binding, or use of streptavidin/biotin binding wherein the immobilized compound includes a biotin moiety. Detection of binding can be accomplished (i) using a radioactive label on the compound that is not immobilized, (ii) using of a fluorescent label on the non-immobilized compound, (iii) using an antibody immunospecific for the non-immobilized compound, (iv) using a label on the non-immobilized compound that excites a fluorescent support to which the immobilized compound is attached, as well as other techniques well known and routinely practiced in the art.

The invention also provides cell-based assays to identify binding partner compounds of a Con-218 polypeptide. In one embodiment, the invention provides a method comprising the steps of contacting a Con-218 polypeptide expressed on the surface of a cell with a candidate binding partner compound and detecting binding of the candidate binding partner compound to the Con-218 polypeptide. In a preferred embodiment, the detection comprises detecting a calcium flux or other physiological event in the cell caused by the binding of the molecule.

Another aspect of the present invention is directed to methods of identifying compounds that bind to either Con-218 or nucleic acid molecules encoding Con-218, comprising contacting Con-218, or a nucleic acid molecule encoding the same, with a compound, and determining

whether the compound binds Con-218 or a nucleic acid molecule encoding the same. Binding can be determined by binding assays which are well known to the skilled artisan, including, but not limited to, gel-shift assays, Western blots, radiolabeled competition assay, phage-based expression cloning, co-fractionation by chromatography, co-precipitation, cross linking, interaction trap/two-hybrid analysis, southwestern analysis, ELISA, and the like, which are described in, for example, *Current Protocols in Molecular Biology*, 1999, John Wiley & Sons, NY, which is incorporated herein by reference in its entirety. The compounds to be screened include (which may include compounds which are suspected to bind Con-218, or a nucleic acid molecule encoding the same), but are not limited to, extracellular, intracellular, biologic or chemical origin. The methods of the invention also embrace ligands, especially neuropeptides, that are attached to a label, such as a radiolabel (*e.g.*,  $^{125}\text{I}$ ,  $^{35}\text{S}$ ,  $^{32}\text{P}$ ,  $^{33}\text{P}$ ,  $^3\text{H}$ ), a fluorescence label, a chemiluminescent label, an enzymic label and an immunogenic label. Modulators falling within the scope of the invention include, but are not limited to, non-peptide molecules such as non-peptide mimetics, non-peptide allosteric effectors, and peptides. The Con-218 polypeptide or polynucleotide employed in such a test may either be free in solution, attached to a solid support, borne on a cell surface or located intracellularly or associated with a portion of a cell. One skilled in the art can, for example, measure the formation of complexes between Con-218 and the compound being tested. Alternatively, one skilled in the art can examine the diminution in complex formation between Con-218 and its substrate caused by the compound being tested.

In another embodiment of the invention, high throughput screening for compounds having suitable binding affinity to Con-218 is employed. Briefly, large numbers of different small peptide test compounds are synthesized on a solid substrate. The peptide test compounds are contacted with Con-218 and washed. Bound Con-218 is then detected by methods well known in the art. Purified polypeptides of the invention can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the protein and immobilize it on the solid support.

Generally, an expressed Con-218 can be used for HTS binding assays in conjunction with its defined ligand, in this case the corresponding neuropeptide that activates it. The identified peptide is labeled with a suitable radioisotope, including, but not limited to,  $^{125}\text{I}$ ,  $^3\text{H}$ ,  $^{35}\text{S}$  or  $^{32}\text{P}$ , by methods that are well known to those skilled in the art. Alternatively, the peptides may be labeled by well-known methods with a suitable fluorescent derivative (Baindur *et al.*, *Drug Dev. Res.*, **1994**, *33*, 373-398; Rogers, *Drug Discovery Today*, **1997**, *2*, 156-160). Radioactive ligand specifically bound to the receptor in membrane preparations made from the cell line expressing the recombinant protein can be detected in HTS assays in one of several standard ways, including filtration of the receptor-ligand complex to separate bound ligand from unbound ligand (Williams, *Med. Res. Rev.*, **1991**, *11*, 147-184; Sweetnam *et al.*, *J. Natural Products*, **1993**, *56*, 441-455). Alternative methods include a scintillation proximity assay (SPA) or a FlashPlate format in which such separation is unnecessary (Nakayama, *Cur. Opinion Drug Disc. Dev.*, **1998**, *1*, 85-91 Bossé *et al.*, *J. Biomolecular Screening*, **1998**, *3*, 285-292.). Binding of fluorescent ligands can be detected in various ways, including fluorescence energy transfer (FRET), direct spectrophotofluorometric analysis of bound ligand, or fluorescence polarization (Rogers, *Drug Discovery Today*, **1997**, *2*, 156-160; Hill, *Cur. Opinion Drug Disc. Dev.*, **1998**, *1*, 92-97).

Other assays may be used to identify specific ligands of a Con-218 receptor, including assays that identify ligands of the target protein through measuring direct binding of test ligands to the target protein, as well as assays that identify ligands of target proteins through affinity ultrafiltration with ion spray mass spectroscopy/HPLC methods or other physical and analytical methods. Alternatively, such binding interactions are evaluated indirectly using the yeast two-hybrid system described in Fields *et al.*, *Nature*, 340:245-246 (1989), and Fields *et al.*, *Trends in Genetics*, 10:286-292 (1994), both of which are incorporated herein by reference. The two-hybrid system is a genetic assay for detecting interactions between two proteins or polypeptides. It can be used to identify proteins that bind to a known protein of interest, or to delineate domains or residues critical for an interaction. Variations on this methodology have been

developed to clone genes that encode DNA binding proteins, to identify peptides that bind to a protein, and to screen for drugs. The two-hybrid system exploits the ability of a pair of interacting proteins to bring a transcription activation domain into close proximity with a DNA binding domain that binds to an upstream activation sequence (UAS) of a reporter gene, and is generally performed in yeast. The assay requires the construction of two hybrid genes encoding (1) a DNA-binding domain that is fused to a first protein and (2) an activation domain fused to a second protein. The DNA-binding domain targets the first hybrid protein to the UAS of the reporter gene; however, because most proteins lack an activation domain, this DNA-binding hybrid protein does not activate transcription of the reporter gene. The second hybrid protein, which contains the activation domain, cannot by itself activate expression of the reporter gene because it does not bind the UAS. However, when both hybrid proteins are present, the noncovalent interaction of the first and second proteins tethers the activation domain to the UAS, activating transcription of the reporter gene. For example, when the first protein is a GPCR gene product, or fragment thereof, that is known to interact with another protein or nucleic acid, this assay can be used to detect agents that interfere with the binding interaction. Expression of the reporter gene is monitored as different test agents are added to the system. The presence of an inhibitory agent results in lack of a reporter signal.

The function of Con-218 gene products is unclear and no ligands have yet been found which bind the gene product. The yeast two-hybrid assay can also be used to identify proteins that bind to the gene product. In an assay to identify proteins that bind to a Con-218 receptor, or fragment thereof, a fusion polynucleotide encoding both a Con-218 receptor (or fragment) and a UAS binding domain (*i.e.*, a first protein) may be used. In addition, a large number of hybrid genes each encoding a different second protein fused to an activation domain are produced and screened in the assay. Typically, the second protein is encoded by one or more members of a total cDNA or genomic DNA fusion library, with each second protein-coding region being fused to the activation domain. This system is applicable to a wide variety of proteins, and it is not even necessary to know the identity or function of the second binding protein. The system is



highly sensitive and can detect interactions not revealed by other methods; even transient interactions may trigger transcription to produce a stable mRNA that can be repeatedly translated to yield the reporter protein.

Other assays may be used to search for agents that bind to the target protein. One such screening method to identify direct binding of test ligands to a target protein is described in U.S. Patent No. 5,585,277, incorporated herein by reference. This method relies on the principle that proteins generally exist as a mixture of folded and unfolded states, and continually alternate between the two states. When a test ligand binds to the folded form of a target protein (*i.e.*, when the test ligand is a ligand of the target protein), the target protein molecule bound by the ligand remains in its folded state. Thus, the folded target protein is present to a greater extent in the presence of a test ligand which binds the target protein, than in the absence of a ligand. Binding of the ligand to the target protein can be determined by any method that distinguishes between the folded and unfolded states of the target protein. The function of the target protein need not be known in order for this assay to be performed. Virtually any agent can be assessed by this method as a test ligand, including, but not limited to, metals, polypeptides, proteins, lipids, polysaccharides, polynucleotides and small organic molecules.

Another method for identifying ligands of a target protein is described in Wieboldt *et al.*, Anal. Chem., 69:1683-1691 (1997), incorporated herein by reference. This technique screens combinatorial libraries of 20-30 agents at a time in solution phase for binding to the target protein. Agents that bind to the target protein are separated from other library components by simple membrane washing. The specifically selected molecules that are retained on the filter are subsequently liberated from the target protein and analyzed by HPLC and pneumatically assisted electrospray (ion spray) ionization mass spectroscopy. This procedure selects library components with the greatest affinity for the target protein, and is particularly useful for small molecule libraries.

Other embodiments of the invention comprise using competitive screening assays in which neutralizing antibodies capable of binding a polypeptide of the invention specifically

compete with a test compound for binding to the polypeptide. In this manner, the antibodies can be used to detect the presence of any peptide that shares one or more antigenic determinants with Con-218. Radiolabeled competitive binding studies are described in A.H. Lin *et al.*

*Antimicrobial Agents and Chemotherapy*, 1997, vol. 41, no. 10. pp. 2127-2131, the disclosure of which is incorporated herein by reference in its entirety.

As described above and in Example 4 below, the gene encoding Con-218 has been detected in brain tissue indicating that these Con-218 proteins are neuroreceptors. Accordingly, natural binding partners of these molecules include neurotransmitters.

#### **Identification of modulating agents**

The invention also provides methods for identifying a modulator of binding between a Con-218 and a Con-218 binding partner, comprising the steps of: (a) contacting a Con-218 binding partner and a composition comprising a Con-218 in the presence and in the absence of a putative modulator compound; (b) detecting binding between the binding partner and the Con-218; and (c) identifying a putative modulator compound or a modulator compound in view of decreased or increased binding between the binding partner and the Con-218 in the presence of the putative modulator, as compared to binding in the absence of the putative modulator.

Con-218 binding partners that stimulate Con-218 activity are useful as agonists in disease states or conditions characterized by insufficient Con-218 signaling (*e.g.*, as a result of insufficient activity of a Con-218 ligand). Con-218 binding partners that block ligand-mediated Con-218 signaling are useful as Con-218 antagonists to treat disease states or conditions characterized by excessive Con-218 signaling. In addition Con-218 modulators in general, as well as Con-218 polynucleotides and polypeptides, are useful in diagnostic assays for such diseases or conditions.

In another aspect, the invention provides methods for treating a disease or abnormal condition by administering to a patient in need of such treatment a substance that modulates the activity or expression of a polypeptide having a sequence selected from the group consisting of: SEQ ID NO:2 and SEQ ID NO:4.

Agents that modulate (*i.e.*, increase, decrease, or block) Con-218 activity or expression may be identified by incubating a putative modulator with a cell containing a Con-218 polypeptide or polynucleotide and determining the effect of the putative modulator on Con-218 activity or expression. The selectivity of a compound that modulates the activity of Con-218 can be evaluated by comparing its effects on Con-218 to its effect on other GPCR compounds. Selective modulators may include, for example, antibodies and other proteins, peptides, or organic molecules that specifically bind to a Con-218 polypeptide or a Con-218-encoding nucleic acid. Modulators of Con-218 activity will be therapeutically useful in treatment of diseases and physiological conditions in which normal or aberrant Con-218 activity is involved. Con-218 polynucleotides, polypeptides, and modulators may be used in the treatment of such diseases and conditions as infections, such as viral infections caused by HIV-1 or HIV-2; pain; cancers; Parkinson's disease; hypotension; hypertension; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Tourette's Syndrome, among others. Con-218 polynucleotides and polypeptides, as well as Con-218 modulators, may also be used in diagnostic assays for such diseases or conditions.

Methods of the invention to identify modulators include variations on any of the methods described above to identify binding partner compounds, the variations including techniques wherein a binding partner compound has been identified and the binding assay is carried out in the presence and absence of a candidate modulator. A modulator is identified in those instances where binding between the Con-218 polypeptide and the binding partner compound changes in the presence of the candidate modulator compared to binding in the absence of the candidate modulator compound. A modulator that increases binding between the Con-218 polypeptide and the binding partner compound is described as an enhancer or activator, and a modulator that decreases binding between the Con-218 polypeptide and the binding partner compound is described as an inhibitor.

The invention also comprehends high-throughput screening (HTS) assays to identify compounds that interact with or inhibit biological activity (*i.e.*, affect enzymatic activity, binding activity, *etc.*) of a Con-218 polypeptide. HTS assays permit screening of large numbers of compounds in an efficient manner. Cell-based HTS systems are contemplated to investigate Con-218 receptor-ligand interaction. HTS assays are designed to identify “hits” or “lead compounds” having the desired property, from which modifications can be designed to improve the desired property. Chemical modification of the “hit” or “lead compound” is often based on an identifiable structure/activity relationship between the “hit” and the Con-218 polypeptide.

Another aspect of the present invention is directed to methods of identifying compounds which modulate (*i.e.*, increase or decrease) activity of Con-218 comprising contacting Con-218 with a compound, and determining whether the compound modifies activity of Con-218. The activity in the presence of the test compared is measured to the activity in the absence of the test compound. Where the activity of the sample containing the test compound is higher than the activity in the sample lacking the test compound, the compound will have increased activity. Similarly, where the activity of the sample containing the test compound is lower than the activity in the sample lacking the test compound, the compound will have inhibited activity.

The present invention is particularly useful for screening compounds by using Con-218 in any of a variety of drug screening techniques. The compounds to be screened include (which may include compounds which are suspected to modulate Con-218 activity), but are not limited to, extracellular, intracellular, biologic or chemical origin. The Con-218 polypeptide employed in such a test may be in any form, preferably, free in solution, attached to a solid support, borne on a cell surface or located intracellularly. One skilled in the art can, for example, measure the formation of complexes between Con-218 and the compound being tested. Alternatively, one skilled in the art can examine the diminution in complex formation between Con-218 and its substrate caused by the compound being tested.

The activity of Con-218 polypeptides of the invention can be determined by, for example, examining the ability to bind or be activated by chemically synthesized peptide ligands.

Alternatively, the activity of Con-218 polypeptides can be assayed by examining their ability to bind calcium ions, hormones, chemokines, neuropeptides, neurotransmitters, nucleotides, lipids, odorants, and photons. Alternatively, the activity of the Con-218 polypeptides can be determined by examining the activity of effector molecules including, but not limited to, adenylate cyclase, phospholipases and ion channels. Thus, modulators of Con-218 polypeptide activity may alter a GPCR receptor function, such as a binding property of a receptor or an activity such as G protein-mediated signal transduction or membrane localization. In various embodiments of the method, the assay may take the form of an ion flux assay, a yeast growth assay, a non-hydrolyzable GTP assay such as a [<sup>35</sup>S]-GTP S assay, a cAMP assay, an inositol triphosphate assay, a diacylglycerol assay, an Aequorin assay, a Luciferase assay, a FLIPR assay for intracellular Ca<sup>2+</sup> concentration, a mitogenesis assay, a MAP Kinase activity assay, an arachidonic acid release assay (*e.g.*, using [<sup>3</sup>H]-arachidonic acid), and an assay for extracellular acidification rates, as well as other binding or function-based assays of Con-218 activity that are generally known in the art. In several of these embodiments, the invention comprehends the inclusion of any of the G proteins known in the art, such as G<sub>16</sub>, G<sub>15</sub>, or chimeric G<sub>qds</sub>, G<sub>qs5</sub>, G<sub>qo5</sub>, G<sub>q25</sub>, and the like. Con-218 activity can be determined by methodologies that are used to assay for FaRP activity, which is well known to those skilled in the art. Biological activities of Con-218 receptors according to the invention include, but are not limited to, the binding of a natural or an unnatural ligand, as well as any one of the functional activities of GPCRs known in the art. Non-limiting examples of GPCR activities include transmembrane signaling of various forms, which may involve G protein association and/or the exertion of an influence over G protein binding of various guanidylate nucleotides; another exemplary activity of GPCRs is the binding of accessory proteins or polypeptides that differ from known G proteins.

The modulators of the invention exhibit a variety of chemical structures, which can be generally grouped into non-peptide mimetics of natural GPCR receptor ligands, peptide and non-peptide allosteric effectors of GPCR receptors, and peptides that may function as activators or inhibitors (competitive, uncompetitive and non-competitive) (*e.g.*, antibody products) of GPCR

receptors. The invention does not restrict the sources for suitable modulators, which may be obtained from natural sources such as plant, animal or mineral extracts, or non-natural sources such as small molecule libraries, including the products of combinatorial chemical approaches to library construction, and peptide libraries. Examples of peptide modulators of GPCR receptors exhibit the following primary structures: GLGPRPLRFamide, GNSFLRFamide, GGPQGPLRFamide, GPSGPLRFamide, PDVDHVFLRFamide, and pyro-EDVDHVFLRFamide.

Other assays can be used to examine enzymatic activity including, but not limited to, photometric, radiometric, HPLC, electrochemical, and the like, which are described in, for example, *Enzyme Assays: A Practical Approach*, eds. R. Eisinger and M. J. Danson, **1992**, Oxford University Press, which is incorporated herein by reference in its entirety.

The use of cDNAs encoding GPCRs in drug discovery programs is well-known; assays capable of testing thousands of unknown compounds per day in high-throughput screens (HTSs) are thoroughly documented. The literature is replete with examples of the use of radiolabelled ligands in HTS binding assays for drug discovery (see Williams, *Medicinal Research Reviews*, **1991**, *11*, 147-184.; Sweetnam, *et al.*, *J. Natural Products*, **1993**, *56*, 441-455 for review). Recombinant receptors are preferred for binding assay HTS because they allow for better specificity (higher relative purity), provide the ability to generate large amounts of receptor material, and can be used in a broad variety of formats (see Hodgson, *Bio/Technology*, **1992**, *10*, 973-980; each of which is incorporated herein by reference in its entirety).

A variety of heterologous systems is available for functional expression of recombinant receptors that are well known to those skilled in the art. Such systems include bacteria (Strosberg, *et al.*, *Trends in Pharmacological Sciences*, **1992**, *13*, 95-98), yeast (Pausch, *Trends in Biotechnology*, **1997**, *15*, 487-494), several kinds of insect cells (Vanden Broeck, *Int. Rev. Cytology*, **1996**, *164*, 189-268), amphibian cells (Jayawickreme *et al.*, *Current Opinion in Biotechnology*, **1997**, *8*, 629-634) and several mammalian cell lines (CHO, HEK293, COS, etc.; see Gerhardt, *et al.*, *Eur. J. Pharmacology*, **1997**, *334*, 1-23). These examples do not preclude

the use of other possible cell expression systems, including cell lines obtained from nematodes (PCT application WO 98/37177).

In preferred embodiments of the invention, methods of screening for compounds that modulate Con-218 activity comprise contacting test compounds with Con-218 and assaying for the presence of a complex between the compound and Con-218. In such assays, the ligand is typically labeled. After suitable incubation, free ligand is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of the particular compound to bind to Con-218.

It is well known that activation of heterologous receptors expressed in recombinant systems results in a variety of biological responses, which are mediated by G proteins expressed in the host cells. Occupation of a GPCR by an agonist results in exchange of bound GDP for GTP at a binding site on the  $G_\alpha$  subunit; one can use a radioactive, non-hydrolyzable derivative of GTP,  $GTP\gamma[^{35}S]$ , to measure binding of an agonist to the receptor (Sim *et al.*, *Neuroreport*, 1996, 7, 729-733). One can also use this binding to measure the ability of antagonists to bind to the receptor by decreasing binding of  $GTP\gamma[^{35}S]$  in the presence of a known agonist. One could therefore construct a HTS based on  $GTP\gamma[^{35}S]$  binding, though this is not the preferred method.

The G proteins required for functional expression of heterologous GPCRs can be native constituents of the host cell or can be introduced through well-known recombinant technology. The G proteins can be intact or chimeric. Often, a nearly universally competent G protein (e.g.,  $G_{\alpha_{16}}$ ) is used to couple any given receptor to a detectable response pathway. G protein activation results in the stimulation or inhibition of other native proteins, events that can be linked to a measurable response.

Examples of such biological responses include, but are not limited to, the following: the ability to survive in the absence of a limiting nutrient in specifically engineered yeast cells (Pausch, *Trends in Biotechnology*, 1997, 15, 487-494); changes in intracellular  $Ca^{2+}$  concentration as measured by fluorescent dyes (Murphy, *et al.*, *Cur. Opinion Drug Disc. Dev.*, 1998, 1, 192-199). Fluorescence changes can also be used to monitor ligand-induced changes in

membrane potential or intracellular pH; an automated system suitable for HTS has been described for these purposes (Schroeder, *et al.*, *J. Biomolecular Screening*, **1996**, *1*, 75-80). Melanophores prepared from *Xenopus laevis* show a ligand-dependent change in pigment organization in response to heterologous GPCR activation; this response is adaptable to HTS formats (Jayawickreme *et al.*, *Cur. Opinion Biotechnology*, **1997**, *8*, 629-634). Assays are also available for the measurement of common second messengers, including cAMP, phosphoinositides and arachidonic acid, but these are not generally preferred for HTS.

Preferred methods of HTS employing these receptors include permanently transfected CHO cells, in which agonists and antagonists can be identified by the ability to specifically alter the binding of GTP $\gamma$ [<sup>35</sup>S] in membranes prepared from these cells. In another embodiment of the invention, permanently transfected CHO cells could be used for the preparation of membranes which contain significant amounts of the recombinant receptor proteins; these membrane preparations would then be used in receptor binding assays, employing the radiolabelled ligand specific for the particular receptor. Alternatively, a functional assay, such as fluorescent monitoring of ligand-induced changes in internal Ca<sup>2+</sup> concentration or membrane potential in permanently transfected CHO cells containing each of these receptors individually or in combination would be preferred for HTS. Equally preferred would be an alternative type of mammalian cell, such as HEK293 or COS cells, in similar formats. More preferred would be permanently transfected insect cell lines, such as *Drosophila* S2 cells. Even more preferred would be recombinant yeast cells expressing the *Drosophila melanogaster* receptors in HTS formats well known to those skilled in the art (*e.g.*, Pausch, *Trends in Biotechnology*, **1997**, *15*, 487-494).

The invention contemplates a multitude of assays to screen and identify inhibitors of ligand binding to Con-218 receptors. In one example, the Con-218 receptor is immobilized and interaction with a binding partner is assessed in the presence and absence of a candidate modulator such as an inhibitor compound. In another example, interaction between the Con-218 receptor and its binding partner is assessed in a solution assay, both in the presence and absence



of a candidate inhibitor compound. In either assay, an inhibitor is identified as a compound that decreases binding between the Con-218 receptor and its binding partner. Another contemplated assay involves a variation of the dihybrid assay wherein an inhibitor of protein/protein interactions is identified by detection of a positive signal in a transformed or transfected host cell, as described in PCT publication number WO 95/20652, published August 3, 1995.

Candidate modulators contemplated by the invention include compounds selected from libraries of either potential activators or potential inhibitors. There are a number of different libraries used for the identification of small molecule modulators, including: (1) chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules. Chemical libraries consist of random chemical structures, some of which are analogs of known compounds or analogs of compounds that have been identified as "hits" or "leads" in other drug discovery screens, some of which are derived from natural products, and some of which arise from non-directed synthetic organic chemistry. Natural product libraries are collections of microorganisms, animals, plants, or marine organisms which are used to create mixtures for screening by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of plants or marine organisms. Natural product libraries include polyketides, non-ribosomal peptides, and variants (non-naturally occurring) thereof. For a review, see Science 282:63-68 (1998). Combinatorial libraries are composed of large numbers of peptides, oligonucleotides, or organic compounds as a mixture. These libraries are relatively easy to prepare by traditional automated synthesis methods, PCR, cloning, or proprietary synthetic methods. Of particular interest are non-peptide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, Curr. Opin. Biotechnol. 8:701-707 (1997). Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to modulate activity.

Still other candidate inhibitors contemplated by the invention can be designed and include soluble forms of binding partners, as well as such binding partners as chimeric, or fusion, proteins. A "binding partner" as used herein broadly encompasses non-peptide modulators, as well as such peptide modulators as neuropeptides other than natural ligands, antibodies, antibody fragments, and modified compounds comprising antibody domains that are immunospecific for the expression product of the identified Con-218 gene.

The polypeptides of the invention are employed as a research tool for identification, characterization and purification of interacting, regulatory proteins. Appropriate labels are incorporated into the polypeptides of the invention by various methods known in the art and the polypeptides are used to capture interacting molecules. For example, molecules are incubated with the labeled polypeptides, washed to remove unbound polypeptides, and the polypeptide complex is quantified. Data obtained using different concentrations of polypeptide are used to calculate values for the number, affinity, and association of polypeptide with the protein complex.

Labeled polypeptides are also useful as reagents for the purification of molecules with which the polypeptide interacts including, but not limited to, inhibitors. In one embodiment of affinity purification, a polypeptide is covalently coupled to a chromatography column. Cells and their membranes are extracted, and various cellular subcomponents are passed over the column. Molecules bind to the column by virtue of their affinity to the polypeptide. The polypeptide-complex is recovered from the column, dissociated and the recovered molecule is subjected to protein sequencing. This amino acid sequence is then used to identify the captured molecule or to design degenerate oligonucleotides for cloning the corresponding gene from an appropriate cDNA library.

Alternatively, compounds may be identified which exhibit similar properties to the ligand for the Con-218 of the invention, but which are smaller and exhibit a longer half time than the endogenous ligand in a human or animal body. When an organic compound is designed, a molecule according to the invention is used as a "lead" compound. The design of mimetics to

known pharmaceutically active compounds is a well-known approach in the development of pharmaceuticals based on such "lead" compounds. Mimetic design, synthesis and testing are generally used to avoid randomly screening a large number of molecules for a target property. Furthermore, structural data deriving from the analysis of the deduced amino acid sequences encoded by the DNAs of the present invention are useful to design new drugs, more specific and therefore with a higher pharmacological potency.

Comparison of the protein sequence of the present invention with the sequences present in all the available databases showed a significant homology with the transmembrane portion of G protein coupled receptors. Accordingly, computer modeling can be used to develop a putative tertiary structure of the proteins of the invention based on the available information of the transmembrane domain of other proteins. Thus, novel ligands based on the predicted structure of Con-218 can be designed.

In a particular embodiment, the novel molecules identified by the screening methods according to the invention are low molecular weight organic molecules, in which case a composition or pharmaceutical composition can be prepared thereof for oral intake, such as in tablets. The compositions, or pharmaceutical compositions, comprising the nucleic acid molecules, vectors, polypeptides, antibodies and compounds identified by the screening methods described herein, can be prepared for any route of administration including, but not limited to, oral, intravenous, cutaneous, subcutaneous, nasal, intramuscular or intraperitoneal. The nature of the carrier or other ingredients will depend on the specific route of administration and particular embodiment of the invention to be administered. Examples of techniques and protocols that are useful in this context are, *inter alia*, found in Remington's Pharmaceutical Sciences, 16<sup>th</sup> edition, Osol, A (ed.), 1980, which is incorporated herein by reference in its entirety.

The dosage of these low molecular weight compounds will depend on the disease state or condition to be treated and other clinical factors such as weight and condition of the human or animal and the route of administration of the compound. For treating human or animals, between approximately 0.5 mg/kg of body weight to 500 mg/kg of body weight of the compound can be

administered. Therapy is typically administered at lower dosages and is continued until the desired therapeutic outcome is observed.

The present compounds and methods, including nucleic acid molecules, polypeptides, antibodies, compounds identified by the screening methods described herein, have a variety of pharmaceutical applications and may be used, for example, to treat or prevent unregulated cellular growth, such as cancer cell and tumor growth. In a particular embodiment, the present molecules are used in gene therapy. For a review of gene therapy procedures, see *e.g.* Anderson, *Science*, **1992**, 256, 808-813, which is incorporated herein by reference in its entirety.

The present invention also encompasses a method of agonizing (stimulating) or antagonizing a Con-218 natural binding partner associated activity in a mammal comprising administering to said mammal an agonist or antagonist to one of the above disclosed polypeptides in an amount sufficient to effect said agonism or antagonism. One embodiment of the present invention, then, is a method of treating diseases in a mammal with an agonist or antagonist of the protein of the present invention comprises administering the agonist or antagonist to a mammal in an amount sufficient to agonize or antagonize Con-218-associated functions.

In an effort to discover novel treatments for diseases, biomedical researchers and chemists have designed, synthesized, and tested molecules that inhibit the function of protein polypeptides. Some small organic molecules form a class of compounds that modulate the function of protein polypeptides. Examples of molecules that have been reported to inhibit the function of protein kinases include, but are not limited to, bis monocyclic, bicyclic or heterocyclic aryl compounds (PCT WO 92/20642, published November 26, 1992 by Maguire *et al.*), vinylene-azaindole derivatives (PCT WO 94/14808, published July 7, 1994 by Ballinari *et al.*), 1-cyclopropyl-4-pyridyl-quinolones (U.S. Patent No. 5,330,992), styryl compounds (U.S. Patent No. 5,217,999), styryl-substituted pyridyl compounds (U.S. Patent No. 5,302,606), certain quinazoline derivatives (EP Application No. 0 566 266 A1), seleoindoles and selenides (PCT WO 94/03427, published February 17, 1994 by Denny *et al.*), tricyclic polyhydroxylic

compounds (PCT WO 92/21660, published December 10, 1992 by Dow), and benzylphosphonic acid compounds (PCT WO 91/15495, published October 17, 1991 by Dow *et al*), all of which are incorporated by reference herein, including any drawings.

Exemplary diseases and conditions amenable to treatment based on the present invention include, but are not limited to, thyroid disorders (*e.g.* thyreotoxicosis, myxoedema); renal failure; inflammatory conditions (*e.g.*, Chron's disease); diseases related to cell differentiation and homeostasis; rheumatoid arthritis; autoimmune disorders; movement disorders; CNS disorders (*e.g.*, pain including migraine; stroke; psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, anxiety, generalized anxiety disorder, post-traumatic-stress disorder, depression, bipolar disorder, delirium, dementia, severe mental retardation; dyskinesias, such as Huntington's disease or Tourette's Syndrome; attention disorders including ADD and ADHD, and degenerative disorders such as Parkinson's, Alzheimer's; movement disorders, including ataxias, supranuclear palsy, *etc.*); infections, such as viral infections caused by HIV-1 or HIV-2; metabolic and cardiovascular diseases and disorders (*e.g.*, type 2 diabetes, obesity, anorexia, hypotension, hypertension, thrombosis, myocardial infarction, cardiomyopathies, atherosclerosis, *etc.*); proliferative diseases and cancers (*e.g.*, different cancers such as breast, colon, lung, *etc.*, and hyperproliferative disorders such as psoriasis, prostate hyperplasia, *etc.*); hormonal disorders (*e.g.*, male/female hormonal replacement, polycystic ovarian syndrome, alopecia, *etc.*); sexual dysfunction, among others.

Compounds that can traverse cell membranes and are resistant to acid hydrolysis are potentially advantageous as therapeutics as they can become highly bioavailable after being administered orally to patients. However, many of these protein inhibitors only weakly inhibit function. In addition, many inhibit a variety of protein kinases and will therefore cause multiple side effects as therapeutics for diseases.

Some indolinone compounds, however, form classes of acid resistant and membrane permeable organic molecules. WO 96/22976 (published August 1, 1996 by Ballinari *et al.*) describes hydrosoluble indolinone compounds that harbor tetralin, naphthalene, quinoline, and

indole substituents fused to the oxindole ring. These bicyclic substituents are in turn substituted with polar groups including hydroxylated alkyl, phosphate, and ether substituents. U.S. Patent Application Serial Nos. 08/702,232, filed August 23, 1996, entitled "Indolinone Combinatorial Libraries and Related Products and Methods for the Treatment of Disease" by Tang *et al.* (Lyon & Lyon Docket No. 221/187) and 08/485,323, filed June 7, 1995, entitled "Benzylidene-Z-Indoline Compounds for the Treatment of Disease" by Tang *et al.* (Lyon & Lyon Docket No. 223/298) and International Patent Publication WO 96/22976, published August 1, 1996 by Ballinari *et al.*, all of which are incorporated herein by reference in their entirety, including any drawings, describe indolinone chemical libraries of indolinone compounds harboring other bicyclic moieties as well as monocyclic moieties fused to the oxindole ring. Applications 08/702,232, filed August 23, 1996, entitled "Indolinone Combinatorial Libraries and Related Products and Methods for the Treatment of Disease" by Tang *et al.* (Lyon & Lyon Docket No. 221/187), 08/485,323, filed June 7, 1995, entitled "Benzylidene-Z-Indoline Compounds for the Treatment of Disease" by Tang *et al.* (Lyon & Lyon Docket No. 223/298), and WO 96/22976, published August 1, 1996 by Ballinari *et al.* teach methods of indolinone synthesis, methods of testing the biological activity of indolinone compounds in cells, and inhibition patterns of indolinone derivatives, both of which are incorporated by reference herein, including any drawings.

Other examples of substances capable of modulating kinase activity include, but are not limited to, tyrphostins, quinazolines, quinoxolines, and quinolines. The quinazolines, tyrphostins, quinolines, and quinoxolines referred to above include well-known compounds such as those described in the literature. For example, representative publications describing quinazolines include Barker *et al.*, EPO Publication No. 0 520 722 A1; Jones *et al.*, U.S. Patent No. 4,447,608; Kabbe *et al.*, U.S. Patent No. 4,757,072; Kaul and Vougioukas, U.S. Patent No. 5, 316,553; Kreighbaum and Comer, U.S. Patent No. 4,343,940; Pegg and Wardleworth, EPO Publication No. 0 562 734 A1; Barker *et al.*, Proc. of Am. Assoc. for Cancer Research 32:327 (1991); Bertino, J.R., Cancer Research 3:293-304 (1979); Bertino, J.R., Cancer Research 9(2 part

1):293-304 (1979); Curtin *et al.*, Br. J. Cancer 53:361-368 (1986); Fernandes *et al.*, Cancer Research 43:1117-1123 (1983); Ferris *et al.* J. Org. Chem. 44(2):173-178; Fry *et al.*, Science 265:1093-1095 (1994); Jackman *et al.*, Cancer Research 51:5579-5586 (1981); Jones *et al.* J. Med. Chem. 29(6):1114-1118; Lee and Skibo, Biochemistry 26(23):7355-7362 (1987); Lemus *et al.*, J. Org. Chem. 54:3511-3518 (1989); Ley and Seng, Synthesis 1975:415-522 (1975); Maxwell *et al.*, Magnetic Resonance in Medicine 17:189-196 (1991); Mini *et al.*, Cancer Research 45:325-330 (1985); Phillips and Castle, J. Heterocyclic Chem. 17(19):1489-1596 (1980); Reece *et al.*, Cancer Research 47(11):2996-2999 (1977); Sculier *et al.*, Cancer Immunol. and Immunother. 23:A65 (1986); Sikora *et al.*, Cancer Letters 23:289-295 (1984); and Sikora *et al.*, Analytical Biochem. 172:344-355 (1988), all of which are incorporated herein by reference in their entirety, including any drawings.

Quinoxaline is described in Kaul and Vougioukas, U.S. Patent No. 5,316,553, incorporated herein by reference in its entirety, including any drawings.

Quinolines are described in Dolle *et al.*, J. Med. Chem. 37:2627-2629 (1994); MaGuire, J. Med. Chem. 37:2129-2131 (1994); Burke *et al.*, J. Med. Chem. 36:425-432 (1993); and Burke *et al.* BioOrganic Med. Chem. Letters 2:1771-1774 (1992), all of which are incorporated by reference in their entirety, including any drawings.

Tyrphostins are described in Allen *et al.*, Clin. Exp. Immunol. 91:141-156 (1993); Anafi *et al.*, Blood 82:12:3524-3529 (1993); Baker *et al.*, J. Cell Sci. 102:543-555 (1992); Bilder *et al.*, Amer. Physiol. Soc. pp. 6363-6143:C721-C730 (1991); Brunton *et al.*, Proceedings of Amer. Assoc. Cancer Rsch. 33:558 (1992); Bryckaert *et al.*, Experimental Cell Research 199:255-261 (1992); Dong *et al.*, J. Leukocyte Biology 53:53-60 (1993); Dong *et al.*, J. Immunol. 151(5):2717-2724 (1993); Gazit *et al.*, J. Med. Chem. 32:2344-2352 (1989); Gazit *et al.*, J. Med. Chem. 36:3556-3564 (1993); Kaur *et al.*, Anti-Cancer Drugs 5:213-222 (1994); King *et al.*, Biochem. J. 275:413-418 (1991); Kuo *et al.*, Cancer Letters 74:197-202 (1993); Levitzki, A., The FASEB J. 6:3275-3282 (1992); Lyall *et al.*, J. Biol. Chem. 264:14503-14509 (1989); Peterson *et al.*, The Prostate 22:335-345 (1993); Pillemer *et al.*, Int. J. Cancer 50:80-85 (1992);

Posner *et al.*, Molecular Pharmacology 45:673-683 (1993); Rendu *et al.*, Biol. Pharmacology 44(5):881-888 (1992); Sauro and Thomas, Life Sciences 53:371-376 (1993); Sauro and Thomas, J. Pharm. and Experimental Therapeutics 267(3):119-1125 (1993); Wolbring *et al.*, J. Biol. Chem. 269(36):22470-22472 (1994); and Yoneda *et al.*, Cancer Research 51:4430-4435 (1991); all of which are incorporated herein by reference in their entirety, including any drawings.

Other compounds that could be used as modulators include oxindolinones such as those described in U.S. patent application Serial No. 08/702,232 filed August 23, 1996, incorporated herein by reference in its entirety, including any drawings.

Methods of determining the dosages of compounds to be administered to a patient and modes of administering compounds to an organism are disclosed in U.S. Application Serial No. 08/702,282, filed August 23, 1996 and International patent publication number WO 96/22976, published August 1 1996, both of which are incorporated herein by reference in their entirety, including any drawings, figures or tables. Those skilled in the art will appreciate that such descriptions are applicable to the present invention and can be easily adapted to it.

The proper dosage depends on various factors such as the type of disease being treated, the particular composition being used and the size and physiological condition of the patient. Therapeutically effective doses for the compounds described herein can be estimated initially from cell culture and animal models. For example, a dose can be formulated in animal models to achieve a circulating concentration range that initially takes into account the  $IC_{50}$  as determined in cell culture assays. The animal model data can be used to more accurately determine useful doses in humans.

Plasma half-life and biodistribution of the drug and metabolites in the plasma, tumors and major organs can also be determined to facilitate the selection of drugs most appropriate to inhibit a disorder. Such measurements can be carried out. For example, HPLC analysis can be performed on the plasma of animals treated with the drug and the location of radiolabeled compounds can be determined using detection methods such as X-ray, CAT scan and MRI. Compounds that show potent inhibitory activity in the screening assays, but have poor pharm-



acokinetic characteristics, can be optimized by altering the chemical structure and retesting. In this regard, compounds displaying good pharmacokinetic characteristics can be used as a model.

Toxicity studies can also be carried out by measuring the blood cell composition. For example, toxicity studies can be carried out in a suitable animal model as follows: 1) the compound is administered to mice (an untreated control mouse should also be used); 2) blood samples are periodically obtained via the tail vein from one mouse in each treatment group; and 3) the samples are analyzed for red and white blood cell counts, blood cell composition and the percent of lymphocytes versus polymorphonuclear cells. A comparison of results for each dosing regime with the controls indicates if toxicity is present.

At the termination of each toxicity study, further studies can be carried out by sacrificing the animals (preferably, in accordance with the American Veterinary Medical Association guidelines Report of the American Veterinary Medical Assoc. Panel on Euthanasia, Journal of American Veterinary Medical Assoc., 202:229-249, 1993). Representative animals from each treatment group can then be examined by gross necropsy for immediate evidence of metastasis, unusual illness or toxicity. Gross abnormalities in tissue are noted and tissues are examined histologically. Compounds causing a reduction in body weight or blood components are less preferred, as are compounds having an adverse effect on major organs. In general, the greater the adverse effect the less preferred the compound.

For the treatment of cancers the expected daily dose of a hydrophobic pharmaceutical agent is between 1 to 500 mg/day, preferably 1 to 250 mg/day, and most preferably 1 to 50 mg/day. Drugs can be delivered less frequently provided plasma levels of the active moiety are sufficient to maintain therapeutic effectiveness. Plasma levels should reflect the potency of the drug. Generally, the more potent the compound the lower the plasma levels necessary to achieve efficacy.

Con-218 mRNA transcripts have been found in brain and testis. Within the brain, the mRNA is abundant in several areas including, but not limited to, the hypothalamus, medial habenular nucleus, hippocampus, and piriform cortex. SEQ ID NOs: 1 and/or 3 will, as detailed

above, enable screening the endogenous neurotransmitters/hormones/ligands which activate, agonize, or antagonize Con-218 and for compounds with potential utility in treating disorders including, but not limited to, CNS disorders (e.g., pain including migraine; stroke; psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, anxiety, generalized anxiety disorder, post-traumatic-stress disorder, depression, bipolar disorder, delirium, dementia, severe mental retardation; dyskinesias, such as Huntington's disease or Tourette's Syndrome; attention disorders including ADD and ADHD, and degenerative disorders such as Parkinson's, Alzheimer's; movement disorders, including ataxias, supranuclear palsy, etc.); infections, such as viral infections caused by HIV-1 or HIV-2; metabolic and cardiovascular diseases and disorders (e.g., type 2 diabetes, obesity, anorexia, hypotension, hypertension, thrombosis, myocardial infarction, cardiomyopathies, atherosclerosis, etc.); hormonal disorders (e.g., male/female hormonal replacement, polycystic ovarian syndrome, alopecia, etc.); sexual dysfunction, among others.

For example, Con-218 receptor activation may mediate neuronal and astrocyte apoptosis and prevention of neurite outgrowth. Inhibition would be beneficial in both chronic and acute brain injury. See, e.g., Donovan *et al.* (1997) J Neurosci 17:5316-5326; Turgeon *et al.* (1998) J Neurosci 18:6882-6891; Smith-Swintosky *et al.* (1997) J Neurochem 69:1890-1896; Gill *et al.* (1998) Brain Res 797:321-327; Suidan *et al.* (1996) Semin Thromb Hemost 22:125-133.

The attached Sequence Listing contains the sequences of the polynucleotides and polypeptides of the invention and is incorporated herein by reference in its entirety.

As described above and in Example 5 below, the gene encoding Con-218 has been detected in brain tissue indicating that these Con-218 proteins are neuroreceptors. The identification of modulators such as agonists and antagonists is therefore useful for the identification of compounds useful to treat neurological diseases and disorders. Such neurological diseases and disorders, including but are not limited to, schizophrenia, affective disorders, ADHD/ADD (*i.e.*, Attention Deficit-Hyperactivity Disorder/Attention Deficit Disorder), and neural disorders such as Alzheimer's disease, Parkinson's disease, migraine, and

senile dementia as well as depression, anxiety, bipolar disease, epilepsy, neuritis, neurasthenia, neuropathy, neuroses, and the like. Further, modulators of Con-218 activity have utility for treating metabolic disorders (including obesity and diabetes), endocrine disorders, as well as inflammatory disorders and cancers. Use of Con-218 modulators, including Con-218 ligands and anti-Con-218 antibodies, to treat individuals having such disease states is intended as an aspect of the invention.

### **Methods of Screening Human Subjects**

Thus in yet another embodiment, the invention provides genetic screening procedures that entail analyzing a person's genome -- in particular their alleles for GPCRs of the invention -- to determine whether the individual possesses a genetic characteristic found in other individuals that are considered to be afflicted with, or at risk for, developing a mental disorder or disease of the brain that is suspected of having a hereditary component. For example, in one embodiment, the invention provides a method for determining a potential for developing a disorder affecting the brain in a human subject comprising the steps of analyzing the coding sequence of one or more GPCR genes from the human subject; and determining development potential for the disorder in said human subject from the analyzing step.

More particularly, the invention provides a method of screening a human subject to diagnose a disorder affecting the brain or genetic predisposition therefor, comprising the steps of: (a) assaying nucleic acid of a human subject to determine a presence or an absence of a mutation altering the amino acid sequence, expression, or biological activity of at least one seven transmembrane receptor that is expressed in the brain, wherein the seven transmembrane receptor comprises an amino acid sequence selected from the group consisting of SEQ ID NO:2 or SEQ ID NO:4, or an allelic variant thereof, and wherein the nucleic acid corresponds to the gene encoding the seven transmembrane receptor; and (b) diagnosing the disorder or predisposition from the presence or absence of said mutation, wherein the presence of a mutation altering the amino acid sequence, expression, or biological activity of allele in the nucleic acid correlates with an increased risk of developing the disorder.

By “human subject” is meant any human being, human embryo, or human fetus. It will be apparent that methods of the present invention will be of particular interest to individuals that have themselves been diagnosed with a disorder affecting the brain or have relatives that have been diagnosed with a disorder affecting the brain.

By “screening for an increased risk” is meant determination of whether a genetic variation exists in the human subject that correlates with a greater likelihood of developing a disorder affecting the brain than exists for the human population as a whole, or for a relevant racial or ethnic human sub-population to which the individual belongs. Both positive and negative determinations (i.e., determinations that a genetic predisposition marker is present or is absent) are intended to fall within the scope of screening methods of the invention. In preferred embodiments, the presence of a mutation altering the sequence or expression of at least one Con-218 seven transmembrane receptor allele in the nucleic acid is correlated with an increased risk of developing schizophrenia, whereas the absence of such a mutation is reported as a negative determination.

The “assaying” step of the invention may involve any techniques available for analyzing nucleic acid to determine its characteristics, including but not limited to well-known techniques such as single-strand conformation polymorphism analysis (SSCP) [Orita *et al.*, *Proc Natl. Acad. Sci. USA*, 86: 2766-2770 (1989)]; heteroduplex analysis [White *et al.*, *Genomics*, 12: 301-306 (1992)]; denaturing gradient gel electrophoresis analysis [Fischer *et al.*, *Proc. Natl. Acad. Sci. USA*, 80: 1579-1583 (1983); and Riesner *et al.*, *Electrophoresis*, 10: 377-389 (1989)]; DNA sequencing; RNase cleavage [Myers *et al.*, *Science*, 230: 1242-1246 (1985)]; chemical cleavage of mismatch techniques [Rowley *et al.*, *Genomics*, 30: 574-582 (1995); and Roberts *et al.*, *Nucl. Acids Res.*, 25: 3377-3378 (1997)]; restriction fragment length polymorphism analysis; single nucleotide primer extension analysis [Shumaker *et al.*, *Hum. Mutat.*, 7: 346-354 (1996); and Pastinen *et al.*, *Genome Res.*, 7: 606-614 (1997)]; 5’ nuclease assays [Pease *et al.*, *Proc. Natl. Acad. Sci. USA*, 91:5022-5026 (1994)]; DNA Microchip analysis [Ramsay, G., *Nature Biotechnology*, 16: 40-48 (1999); and Chee *et al.*, U.S. Patent No. 5,837,832]; and ligase chain

reaction [Whiteley *et al.*, U.S. Patent No. 5,521,065]. [See generally, Schafer and Hawkins, *Nature Biotechnology*, 16: 33-39 (1998).] All of the foregoing documents are hereby incorporated by reference in their entirety.

Thus, in one preferred embodiment involving screening Con-218 sequences, for example, the assaying step comprises at least one procedure selected from the group consisting of: (a) determining a nucleotide sequence of at least one codon of at least one Con-218 allele of the human subject; (b) performing a hybridization assay to determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences; (c) performing a polynucleotide migration assay to determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences; and (d) performing a restriction endonuclease digestion to determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences.

In a highly preferred embodiment, the assaying involves sequencing of nucleic acid to determine nucleotide sequence thereof, using any available sequencing technique. [See, *e.g.*, Sanger *et al.*, *Proc. Natl. Acad. Sci. (USA)*, 74: 5463-5467 (1977) (dideoxy chain termination method); Mirzabekov, *TIBTECH*, 12: 27-32 (1994) (sequencing by hybridization); Drmanac *et al.*, *Nature Biotechnology*, 16: 54-58 (1998); U.S. Patent No. 5,202,231; and *Science*, 260: 1649-1652 (1993) (sequencing by hybridization); Kieleczawa *et al.*, *Science*, 258: 1787-1791 (1992) (sequencing by primer walking); (Douglas *et al.*, *Biotechniques*, 14: 824-828 (1993) (Direct sequencing of PCR products); and Akane *et al.*, *Biotechniques* 16: 238-241 (1994); Maxam and Gilbert, *Meth. Enzymol.*, 65: 499-560 (1977) (chemical termination sequencing), all incorporated herein by reference.] The analysis may entail sequencing of the entire nGPCR gene genomic DNA sequence, or portions thereof; or sequencing of the entire seven transmembrane receptor coding sequence or portions thereof. In some circumstances, the analysis may involve a determination of whether an individual possesses a particular allelic variant, in which case sequencing of only a small portion of nucleic acid -- enough to determine the sequence of a

particular codon characterizing the allelic variant -- is sufficient. This approach is appropriate, for example, when assaying to determine whether one family member inherited the same allelic variant that has been previously characterized for another family member, or, more generally, whether a person's genome contains an allelic variant that has been previously characterized and correlated with a mental disorder having a heritable component.

In another highly preferred embodiment, the assaying step comprises performing a hybridization assay to determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences. In a preferred embodiment, the hybridization involves a determination of whether nucleic acid derived from the human subject will hybridize with one or more oligonucleotides, wherein the oligonucleotides have nucleotide sequences that correspond identically to a portion of the Con-218 gene sequence taught herein, set forth in SEQ ID NO:1 or SEQ ID NO:3, or that correspond identically except for one mismatch. The hybridization conditions are selected to differentiate between perfect sequence complementarity and imperfect matches differing by one or more bases. Such hybridization experiments thereby can provide single nucleotide polymorphism sequence information about the nucleic acid from the human subject, by virtue of knowing the sequences of the oligonucleotides used in the experiments.

Several of the techniques outlined above involve an analysis wherein one performs a polynucleotide migration assay, *e.g.*, on a polyacrylamide electrophoresis gel (or in a capillary electrophoresis system), under denaturing or non-denaturing conditions. Nucleic acid derived from the human subject is subjected to gel electrophoresis, usually adjacent to (or co-loaded with) one or more reference nucleic acids, such as reference GPCR-encoding sequences having a coding sequence identical to all or a portion of SEQ ID NO:1 or SEQ ID NO:3 (or identical except for one known polymorphism). The nucleic acid from the human subject and the reference sequence(s) are subjected to similar chemical or enzymatic treatments and then electrophoresed under conditions whereby the polynucleotides will show a differential migration pattern, unless they contain identical sequences. [See generally Ausubel *et al.* (eds.), *Current*

*Protocols in Molecular Biology*, New York: John Wiley & Sons, Inc. (1987-1999); and Sambrook *et al.*, (eds.), *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press (1989), both incorporated herein by reference in their entirety.]

In the context of assaying, the term “nucleic acid of a human subject” is intended to include nucleic acid obtained directly from the human subject (*e.g.*, DNA or RNA obtained from a biological sample such as a blood, tissue, or other cell or fluid sample); and also nucleic acid derived from nucleic acid obtained directly from the human subject. By way of non-limiting examples, well known procedures exist for creating cDNA that is complementary to RNA derived from a biological sample from a human subject, and for amplifying (*e.g.*, via polymerase chain reaction (PCR)) DNA or RNA derived from a biological sample obtained from a human subject. Any such derived polynucleotide which retains relevant nucleotide sequence information of the human subject’s own DNA/RNA is intended to fall within the definition of “nucleic acid of a human subject” for the purposes of the present invention.

In the context of assaying, the term “mutation” includes addition, deletion, and/or substitution of one or more nucleotides in the Con-218 gene sequence (*e.g.*, as compared to the seven transmembrane receptor-encoding sequences set forth in SEQ ID NO:1 or SEQ ID NO:3) and other polymorphisms that occur in introns (where introns exist) and that are identifiable via sequencing, restriction fragment length polymorphism, or other techniques. The various activity examples provided herein permit determination of whether a mutation modulates activity of the relevant receptor in the presence or absence of various test substances.

In a related embodiment, the invention provides methods of screening a person’s genotype with respect to the GPCR of the invention, and correlating such genotypes with diagnoses for disease or with predisposition for disease (for genetic counseling). For example, the invention provides a method of screening for an Con-218 hereditary schizophrenia genotype in a human patient, comprising the steps of: (a) providing a biological sample comprising nucleic acid from the patient, the nucleic acid including sequences corresponding to said patient’s

Con-218 alleles; (b) analyzing the nucleic acid for the presence of a mutation or mutations; (c) determining an Con-218 genotype from the analyzing step; and (d) correlating the presence of a mutation in an Con-218 allele with a hereditary schizophrenia genotype. In a preferred embodiment, the biological sample is a cell sample containing human cells that contain genomic DNA of the human subject. The analyzing can be performed analogously to the assaying described in preceding paragraphs. For example, the analyzing comprises sequencing a portion of the nucleic acid (e.g., DNA or RNA), the portion comprising at least one codon of the Con-218.

Although more time consuming and expensive than methods involving nucleic acid analysis, the invention also may be practiced by assaying protein of a human subject to determine the presence or absence of an amino acid sequence variation in GPCR protein from the human subject. Such protein analyses may be performed, *e.g.*, by fragmenting GPCR protein via chemical or enzymatic methods and sequencing the resultant peptides; or by Western analyses using an antibody having specificity for a particular allelic variant of the GPCR.

The invention also provides materials that are useful for performing methods of the invention. For example, the present invention provides oligonucleotides useful as probes in the many analyzing techniques described above. In general, such oligonucleotide probes comprise 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 nucleotides that have a sequence that is identical, or exactly complementary, to a portion of a human GPCR gene sequence taught herein (or allelic variant thereof), or that is identical or exactly complementary except for one nucleotide substitution. In a preferred embodiment, the oligonucleotides have a sequence that corresponds in the foregoing manner to a human GPCR coding sequence taught herein. In one variation, an oligonucleotide probe of the invention is purified and isolated. In another variation, the oligonucleotide probe is labeled, *e.g.*, with a radioisotope, chromophore, or fluorophore. In yet another variation, the probe is covalently attached to a solid support. [See generally Ausubel *et al.* And Sambrook *et al.*, *supra*.]



In a related embodiment, the invention provides kits comprising reagents that are useful for practicing methods of the invention. For example, the invention provides a kit for screening a human subject to diagnose schizophrenia or a genetic predisposition therefor, comprising, in association: (a) an oligonucleotide useful as a probe for identifying polymorphisms in a human Con-218 seven transmembrane receptor gene, the oligonucleotide comprising 6-50 nucleotides that have a sequence that is identical or exactly complementary to a portion of a human Con-218 gene sequence or Con-218 coding sequence, except for one sequence difference selected from the group consisting of a nucleotide addition, a nucleotide deletion, or nucleotide substitution; and (b) a media packaged with the oligonucleotide containing information identifying polymorphisms identifiable with the probe that correlate with schizophrenia or a genetic predisposition therefor. Exemplary information-containing media include printed paper package inserts or packaging labels; and magnetic and optical storage media that are readable by computers or machines used by practitioners who perform genetic screening and counseling services. The practitioner uses the information provided in the media to correlate the results of the analysis with the oligonucleotide with a diagnosis. In a preferred variation, the oligonucleotide is labeled.

In still another embodiment, the invention provides methods of identifying those allelic variants of Con-218 of the invention that correlate with mental disorders. For example, the invention provides a method of identifying a seven transmembrane allelic variant that correlates with a mental disorder, comprising steps of: (a) providing a biological sample comprising nucleic acid from a human patient diagnosed with a mental disorder, or from the patient's genetic progenitors or progeny; (b) analyzing the nucleic acid for the presence of a mutation or mutations in at least one seven transmembrane receptor that is expressed in the brain, wherein the at least one seven transmembrane receptor comprises an amino acid sequence selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:4, or an allelic variant thereof, and wherein the nucleic acid includes sequence corresponding to the gene or genes encoding the at least one seven transmembrane receptor; (c) determining a genotype for the patient for the at least one

seven transmembrane receptor from said analyzing step; and (d) identifying an allelic variant that correlates with the mental disorder from the determining step. To expedite this process, it may be desirable to perform linkage studies in the patients (and possibly their families) to correlate chromosomal markers with disease states.

The foregoing method can be performed to correlate the nGPCR of the invention to a number of disorders having hereditary components that are causative or that predispose persons to the disorder. For example, in one preferred variation, the disorder is schizophrenia, and the at least one seven transmembrane receptor comprises Con-218 having an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO:4, or an allelic variant thereof.

Also contemplated as part of the invention are polynucleotides that comprise the allelic variant sequences identified by such methods, and polypeptides encoded by the allelic variant sequences, and oligonucleotide and oligopeptide fragments thereof that embody the mutations that have been identified. Such materials are useful in *in vitro* cell-free and cell-based assays for identifying lead compounds and therapeutics for treatment of the disorders. For example, the variants are used in activity assays, binding assays, and assays to screen for activity modulators described herein. In one preferred embodiment, the invention provides a purified and isolated polynucleotide comprising a nucleotide sequence encoding a Con-218 receptor allelic variant identified according to the methods described above; and an oligonucleotide that comprises the sequences that differentiate the allelic variant from the Con-218 sequence set forth in SEQ ID NO:1 or SEQ ID NO:3. The invention also provides a vector comprising the polynucleotide (preferably an expression vector); and a host cell transformed or transfected with the polynucleotide or vector. The invention also provides an isolated cell line that is expressing the allelic variant GPCR polypeptide; purified cell membranes from such cells; purified polypeptide; and synthetic peptides that embody the allelic variation amino acid sequence. In one particular embodiment, the invention provides a purified polynucleotide comprising a nucleotide sequence encoding a Con-218 seven transmembrane receptor protein of a human that is affected with schizophrenia; wherein said polynucleotide hybridizes to the complement of SEQ ID NO:1 under

the following hybridization conditions: (a) hybridization for 16 hours at 42°C in a hybridization solution comprising 50% formamide, 1% SDS, 1 M NaCl, 10% dextran sulfate and (b) washing 2 times for 30 minutes at 60°C in a wash solution comprising 0.1x SSC and 1% SDS; and wherein the polynucleotide encodes a Con-218 amino acid sequence that differs from SEQ ID NO: 2 or SEQ ID NO:4 by at least one residue.

An exemplary assay for using the allelic variants is a method for identifying a modulator of Con-218 biological activity, comprising the steps of: (a) contacting a cell expressing the allelic variant in the presence and in the absence of a putative modulator compound; (b) measuring Con-218 biological activity in the cell; and (c) identifying a putative modulator compound in view of decreased or increased Con-218 biological activity in the presence versus absence of the putative modulator.

The following Table 4 contains the sequences of the polynucleotides and polypeptides of the invention. The start and stop codons within the polynucleotide sequence are identified by boldface type. The transmembrane domains within the polypeptide sequence are identified by underlining.

TABLE 4

The following DNA sequence Con-218 <SEQ ID NO. 1> was identified in *H. sapiens*:

AGCCTGGGCTGGGCTCGCCATCCCAGGGTCGCTGGACTAGGATGGGGGATGGGCCTGTGACAGGAGGTA  
CCCTGGGTGCCCTCTTTTCGGCCCC**ATGG**AGTCCTCACCCATCCCCAGTCATCAGGGAACCTCTTCCACTT  
TGGGGAGGGTCCCTCAAACCCCAGGTCCCTCTACTGCCAGTGGGGTCCCGAGGTGGGGCTACGGGATGT  
TGCTTCGGAATCTGTGGCCCTCTTCTTCATGCTCCTGCTGGACTTGACTGCTGTGGCTGGCAATGCCGCT  
GTGATGGCCGTGATCGCCAAGACGCCTGCCCTCCGAAATTTGTCTTCGTCTTCCACCTCTGCCTGGTGG  
ACCTGCTGGCTGCCCTGACCCTCATGCCCCTGGCCATGCTCTCCAGCTCTGCCCTCTTTGACCACGCCCT  
CTTTGGGGAGGTGGCCTGCCGCCTCTACTTGTTTCTGAGCGTGTGCTTTGTGAGCCTGGCCATCCTCTCG  
GTGTGAGCCATCAATGTGGAGCGCTACTATTACGTAGTCCACCCATGCGCTACGAGGTGCGCATGACGC  
TGGGGCTGGTGGCCTCTGTGCTGGTGGGTGTGTGGGTGAAGGCCTTGGCCATGGCTTCTGTGCCAGTGT  
GGGAAGGTCTCCTGGGAGGAAGGAGCTCCAGTGTCCCCCAGGCTGTTCACTCCAGTGGAGCCACAGT  
GCCTACTGCCAGCTTTTGTGGTGGTCTTTGCTGTCTTTACTTTCTGTTGCCCTGCTCCTCATACTTG  
TGGTCTACTGCAGCATGTTCCGAGTGGCCCGCTGGCTGCCATGCAGCACGGGCCGCTGCCACGTGGAT  
GGAGACACCCCGCAACGCTCCGAATCTCTCAGCAGCCGCTCCACGATGGTCACCAGCTCGGGGGCCCCC  
CAGACCACCCACACCGGACGTTTGGGGAGGGAAAGCAGCAGTGGTCTCCTGGCTGTGGGGGACAGT  
TCCTGCTCTGTTGGTTGCCCTACTTCTCTTTCCACCTCTATGTTGCCCTGAGTGTCTAGCCCATTTC AAC  
TGGGCAGGTGGAGAGTGTGGTACCTGGATTGGCTACTTTTGCTTCACTTCCAACCCTTTCTTCTATGGA  
TGTCTCAACCGGCAGATCCGGGGGAGCTCAGCAAGCAGTTTGTCTGCTTCTTCAAGCCAGCTCCAGAGG  
AGGAGCTGAGGCTGCCTAGCCGGGAGGGCTCCATTGAGGAGAACTTCCTGCAGTTCCTTCAGGGGACTGG  
CTGTCTTCTGAGTCTGGGTTTCCCGACCCCTACCCAGCCCCAAGCAGGAGCCACCTGCTGTTGACTTT  
CGAATCCCAGGCCAGATAGCTGAGGAGACCTCTGAGTTCCTGGAGCAGCAACTCACCAGCGACATCATCA  
TGTCAGACAGCTACCTCCGCTCTGCCGCCTCACCCCGGCTGGAGTCA**TGAT**GGGCCGCTGGACACTCGGA  
GGGATATGGGGCTGGGGCCAGTTATGATTGCAAGGACCACCTTGTGGGATCACCTTTTCCAGCTGGCTA

The following amino acid sequence <SEQ ID NO. 2> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 1:

MESSPIPQSSGNSSTLGRVPQTGPGSTASGVPEVGLRDVASESVALFFMLLLDLTAVAGNAAVMAVIAKTP  
ALRKVFVVFHLCVLDLLAALTLMLPLAMLSSSALFDHALFGEVACRLYLFLSVCFVSLAILSVSAINVERYY  
YVVHPMRYEVRMTLGLVASVLGVVVKALAMASVPVLGRVSWEEGAPSVPPGCSLQWSHSAYCQLFVVVFA  
VLYFLLPLLLILVVCYSMFRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMTVSSGAPQTTPHRTFGGK  
AAVLLAVGGQFLLCWLPYFSFHLVALSAQPISTGQVESVVTWIGYFCFTSNPFFYGCLNRQIRGELSKQ  
FVCFKPAPEEELRLPSREGSIEENFLQLQGTGCPSESWSRPLSPKQEPFAVDFRIPGQIAETSEFL  
EQQLTSDIIMSDSYLRPAASPRLES

The following DNA sequence Con-218-RN <SEQ ID NO. 3> was identified in *R. norvegicus*:

GGAAGCCTGGACCTGGGCTTACATCCCAGGGTTGTGGAGTAGGATGGGGGATGGGCCTGTAACAGGAAGTG  
CCCTGGGTGTCTTTTTTCGGCCCC**ATGG**AGTCCTCACCCATCCCCAGTCATCAGGAACTCATCCACTTTG  
GGAAGGGCCCTTCAAACCCCAGGTCCCTCTACGGCCAGCGAGTCCAGAGTTGGGACTAAGGGACGTGGCT  
TCGGAATCTGTGGCCCTCTTCTTCATGCTTCTGTTGGATCTCACTGCTGTGGCGGGCAATGCTGCTGTGATG  
GCTGTTATTGCCAAGACACCCGCCCTCCGAAAGTTTGTCTTCCATCTCTGTCTGGTGGACCTGCTG  
GCTGCCCTGACCCTCATGCCCCTGGCCATGCTCTCCAGCTCTGCCCTCTTTGACCACGCCCTCTTTGGGGAG

GTGGCCTGCCGCTCTACCTGTTCTGAGCGTTTGCTTTGTGAGCCTGGCCatCCTTTCCGGTGTCTGCCATT  
 AATGTGGAGCGCTACTATTATGTGGTCCACCCCATGCGCTACGAGGTGCGCATGACACTAGGGCTGGTGGCC  
 TCTGTGCTGGTGGGCGTGTGGGTAAAGGCCCTGGCCATGGCGTCTGTGCCAGTGTGGGAAGGGTCTACTGG  
 GAGGAAGGAGCTCCAGTGTTAACCAGGCTGTTCTCTCCAATGGAGCCATAGTGCCTACTGCCAGCTTTTT  
 GTGGTGGTCTTTGCTGTTCTTTACTTCTTGCTGCCCTTGATCCTGATCTTTGTGGTCTACTGCAGCATGTTT  
 CGAGTGGCTCGCGTGGCTGCCATGCAACATGGGCCGCTGCCACGTGGATGGAGACGCCCCGGCAACGCTCT  
 GAGTCTCTCAGTAGCCGCTCTACTATGGTCACTAGCTCCGGGGCTCATCAGACCACCCACACCGGACGTTT  
 GGGGGTGGGAAGGCAGCAGTGGTCCTCCTGGCTGTCGGGGGACAGTTCTTGCTTTGTTGGTTACCCTACTTC  
 TCTTTCCATCTCTATGTTGCCCTGAGCGCTCAGCCCATTTCAACAGGACAGGTGGAGAATGTGGTGACCTGG  
 ATCGGCTACTTTTGCTTCACTTCCAACCCTTTTTTCTATGGATGTCTCAACCGTCAGATCCGGGGCGAGCTT  
 AGCAAACAGTTTGTCTGTTTCTTCAAGGCAGCTCCAGAGGAGGAGCTGAGGCTGCCGAGTCGCGAGGGCTCC  
 ATCGAGGAGAATTTCTACAGTTTCTCCAGGTACATCTGAGAACTGGGTTTCTCGGCCCTACCCAGCCCT  
 AAGCGGGAGCCACCGCCTGCTGTTGACTTTTGAATCCCAGGCCAGATTGCTGAGGAGACCTCGGAGTTCTTG  
 GAGCAGCAACTCACCAGCGACATCATGTCGGACAGCTACCTCCGTCCTGCCCTTACCAAGGCTGGAG  
 TCATGATGGACAGACACTAGGAGGGAATAAGGCTTGGGGCTGGTTTATCATCTCAAGGATTGCTTTTCCAGC  
 TGGCTGGGGTTTGGACTCGGGTCTCTGGACTTAGCTTTTGTGTGGTGTTCCTGGGTCAGGACCAGAGTCAA  
 CGGGATGGACATGTGGCAAAAAGCCTTGGACTTGGCTGTGATCTTTGACTATTGGGGGAGGGATCCTGGGTA  
 TGGTGAGACGGTGATGAGAGAAAAGGGTGACAAAGGTGAGGGAAGCCTTTCTACCAGTGAACCTCTTCGTGC  
 CTCAGGAGACAGGGCAACTTCTGGGTAGGCATTGGAGCAGCAGGCTAGGAGCAGTTATTCTGGGGACCGTTG  
 AGGTTTACTTCTTTCCAGTGTCTAGTCCAGACTAATATTTATACTGAGACAAGGTAAGAAAATGGCCACA  
 TCTTCTCATTTGCTAACTAGGTTAAAAAAAAAAAAAAAA

The following amino acid sequence <SEQ ID NO. 4> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 3:

MESSPIQSSGNSSTLGRALQTPGPSTASGVPELGLRDVASESVALFFMLLLDLTAVAGNAAVMAVIAKTPALR  
 KFVFVFLHCLVDLLAALTLMPLAMLSSSALFDHALFGEVACRLYLFLSVCFVSLAILSVSAINVERYYYVVHPM  
 RYEVGMTLGLVASVLGVVVKALAMASVPVLGRVYWEEGAPSVNPGCSLQWSHSAYCQLFVVVFAVLYFLLPLI  
 LIFVVCMSMRVARVAAMQHGPLPTWMETPRQRSESLSSRSTMTSSGAHQTPHRTFGGKAQVLLAVGGQF  
 LLCWLPHYFSFHLVALSAQPISTGQVENVTWIGYFCFTSNPFFYGCLNRQIRGELSKQFVCFKAAPEEELRL  
 PSREGSIEENFLQFLQGTSENWVSRLPSPKREPPPAVDFRIPGQIAEETSEFLEQQLTSDIIMSDSYLRPAPS  
 PRLES

Additional features of the invention will be apparent from the following Examples.

Examples 1, 2, 3, 5, and 6 are actual, while the remaining Examples are prophetic.

### Example 1: Identification of Con-218

The Incyte rat Expressed Sequence Tag database was searched using known GPCR receptors as query sequences to find patterns suggestive of novel G protein-coupled receptors.

Positive hits were further analyzed with the program BLAST to determine which ones were the most likely candidates to encode G protein-coupled receptors, using the standard (default) alignment produced by BLAST as a guide.

Briefly, the BLAST algorithm, which stands for Basic Local Alignment Search Tool is suitable for determining sequence similarity (Altschul *et al.*, J. Molec. Biol., 1990, 215, 403-410, which is incorporated herein by reference in its entirety). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). This algorithm involves first identifying high scoring sequence pair (HSPs) by identifying short words of length W in the query sequence that either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighbourhood word score threshold (Altschul *et al.*, *supra*). These initial neighbourhood word hits act as seeds for initiating searches to find HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Extension for the word hits in each direction are halted when: 1) the cumulative alignment score falls off by the quantity X from its maximum achieved value; 2) the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or 3) the end of either sequence is reached. The Blast algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The Blast program uses as defaults a word length (W) of 11, the BLOSUM62 scoring matrix (see Henikoff *et al.*, Proc. Natl. Acad. Sci. USA, 1992, 89, 10915-10919, which is incorporated herein by reference in its entirety) alignments (B) of 50, expectation (E) of 10, M=5, N=4, and a comparison of both strands.

The BLAST algorithm (Karlin *et al.*, Proc. Natl. Acad. Sci. USA, 1993, 90, 5873-5787, which is incorporated herein by reference) and Gapped BLAST perform a statistical analysis of the similarity between two sequences. One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For

example, a nucleic acid is considered similar to a GPCR gene or cDNA if the smallest sum probability in comparison of the test nucleic acid to a GPCR nucleic acid is less than about 1, preferably less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

Homology searches are performed with the program BLAST version 2.08. A collection of 340 query amino acid sequences derived from GPCR's was used to search the Incyte rat Expressed Sequence Tag database sequences using TBLASTN and alignments with an E-value lower than 0.01 were collected from each BLAST search.

One rat Expressed Sequence Tag (700249079H1) showed significant homology with a putative G protein-coupled receptor isolated from Japanese medaka (*Oryzias latipes*; swissprot: locus GPRX\_ORYLA, accession Q91178) and was subsequently denoted Con-218-RN.

Similar to the above, the Incyte human Expressed Sequence Tag database was queried, and one human sequence (3226980H1) showed significant homology with the same putative G protein-coupled receptor isolated from Japanese medaka (*Oryzias latipes*) and was subsequently denoted Con-218.

### **Example 2: Cloning of Con-218-RN (SEQ ID NO. 3)**

A glycerol stock of the Incyte clone 700249079 with a ~ 1.4 kbp cDNA insert was plated on LB agar with ampicillin (60 micrograms/ml). One colony was picked, cultured over night in 5 ml LB medium with ampicillin (60 microgram/ml) and a small scale plasmid DNA preparation made using Quiaprep Spin Miniprep Kit (Qiagen catalog # 27104). Sequencing was performed using the Dye terminator cycle method with AmpliTaq DNA Polymerase FS on an ABI 377 Sequencer. The clone was shown to contain a 1460 bp insert encoding an open reading frame of 336 amino acids, corresponding to five putative transmembrane regions (transmembrane regions III through VII) as judged by analysis with TMHMM1.0 [Sonnhammer EL, von Heijne G, Krogh A, Ismb 1998;6:175-82], and homology searches using BLAST (see above in Example 1.).

The full length sequence was obtained by performing 5' RACE (*c.f.* Zhang Y, Frohman MA, Methods Mol Biol 1997;69:61-87 ) on rat brain cDNA. Brain polyA<sup>+</sup> RNA was extracted by two repetitive rounds of binding to oligo-dT coated magnetic beads (Dynabeads mRNA Direct kit, Dynal AS, catalog # 61011) from two pooled Sprague-Dawley rat brains. 5'RACE Ready cDNA was synthesized from 1 microgram brain polyA<sup>+</sup> RNA using Superscript II reverse transcriptase as outlined in the SMART RACE cDNA amplification kit protocol (Clontech, catalog # K1811-1). 5'-RACE fragments were amplified by Advantage cDNA Polymerase mix (Clontech, catalog # 8417-1) in reactions with oligonucleotide primers designed based on the sequence information from the 1.4 kbp Incyte clone 700249079 and Universal Primers provided in the kit.

The following oligonucleotide primers were used for the 5'-RACE procedure:

Reverse RACE primer:

5' - TAATGGCAGACACCGAAAGGATGGC - 3' [SEQ ID NO. 5]

Nested reverse RACE primer:

5' - GCTGACAAAGCAAACGCTCAGGAAC - 3' [SEQ ID NO. 6]

The initial 5' RACE cDNA amplification was performed in a 50 microliter reaction according to the SMART RACE cDNA amplification kit protocol in a MJ Research PTC100 Thermocycler using the reverse RACE primer and the following parameters: 94 C for 30 sec followed by five cycles of 72 C for 3 min; 94 C 30 sec followed by 70 C for 30 sec, followed by five cycles of 72 C for 3 min; 94 C for 30 sec, followed by 68 C for 30 sec, followed by 30 cycles of 72 C for 3 min.

One 50th (1 microliter) of this reaction was used as template in a next round of RACE using the nested RACE primer, employing the same cycling parameters. This reaction was subsequently run on an 1.5% agarose gel stained with ethidium bromide, and the main product of



~ 1000 bp was excised and purified using QIAquick Gel Extraction Kit (Qiagen catalog # 28704). The fragments were TA-cloned into pCRII using the TOPO TA cloning kit from InVitrogen (K4600-01) according to the manufacturer's instructions, and transformed into chemically competent *E. coli* Top-10.

96 clones were analyzed by PCR-screening using vector specific primers, Taq DNA Polymerase (Amersham Pharmacia Biotech, catalog # 27-0799-62) and PCR grade dNTPs from Roche (catalog #1 969 064). Three clones were grown overnight in 5 ml LB medium with ampicillin (60 microgram/ml) and subsequently subjected to small scale plasmid DNA preparations using QIAprep Spin Miniprep Kit (Qiagen catalog # 27104). Sequencing was performed with Dye Terminator Cycle Sequencing using AmpliTaq DNA Polymerase on an ABI 377 Sequencer. All three clones were of identical nucleotide sequence, representing the 5'-end of Con-218-RN (SEQ ID NO. 3).

The nucleotide sequences of the 5'-RACE fragments and the initial 1.4 kb Incyte clone were assembled into one continuous nucleotide sequence of 1909 base-pairs (SEQ ID NO. 3) containing an open reading frame encoding 449 amino acids. The translated amino acid sequence (SEQ ID NO. 4) contains seven putative transmembrane regions as judged by TMHMM analysis and BLAST homology searches (see above).

### **Example 3: Detection of human Con-218 (SEQ ID NO. 1)**

The deduced amino acid sequence of Con-218-RN (SEQ ID NO. 4) was used as query sequence to probe the public human genomic nucleotide database in GenBank (<http://www.ncbi.nlm.nih.gov/Genbank/index.html>) using TBLASTN (see Example 1). A high scoring sequence pair was discovered in GenBank AC011780 at around nucleotides 28600 - 31100 which contains the sequence given as SEQ ID NO. 1. Translation of this sequence revealed an open reading frame encoding 451 amino acids (SEQ ID NO.2), which show a high

degree of homology. The previously discovered Incyte human Expressed Sequence Tag 3226980H1 was found to lie within SEQ ID NO. 1 (nucleotides 406 through 602).

#### **Example 4: PCR Screening of Genomic Library and Subcloning of Coding Regions of Con-218**

Two microliters of a human genomic library ( $\sim 10^8$  PFU/ml) (Clontech) are added to 6 ml of an overnight culture of K802 cells (Clontech), then distributed as 250  $\mu$ l aliquots into each of 24 tubes. The tubes are incubated at 37°C for 15 min. Seven milliliters of 0.8% agarose is added to each tube, mixed, then poured onto LB agar + 10 mM  $\text{MgSO}_4$  plates and incubated overnight at 37°C. To each plate 5 ml of SM (0.1M NaCl, 8.1 mM  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , 50mM Tris-Cl (pH 7.5), 0.0001% gelatin) phage buffer is added and the top agarose is removed with a microscope slide and placed in a 50 ml centrifuge tube. A drop of chloroform is added and the tube is placed in a 37 °C shaker for 15 min, then centrifuged for 20 min at 4000 RPM (Sorvall RT6000 table top centrifuge) and the supernatant stored at 4°C as a stock solution.

The PCR reaction may be performed in 20  $\mu$ l containing 8.8  $\mu$ l  $\text{H}_2\text{O}$ , 4  $\mu$ l 5X Rapid-Load Buffer (Origene), 2  $\mu$ l 10xPCR buffer II (Perkin-Elmer), 2  $\mu$ l 25 mM  $\text{MgCl}_2$ , 0.8  $\mu$ l 10 mM dNTP, 0.12  $\mu$ l LW1632 (1  $\mu\text{g}/\mu\text{l}$ ), 0.12  $\mu$ l LW1633 (1  $\mu\text{g}/\mu\text{l}$ ), 0.2  $\mu$ l AmpliTaq Gold polymerase (Perkin Elmer) and 2  $\mu$ l of phage from each of the 24 tubes. The PCR reaction involves 1 cycle at 80°C for 20 min, 95°C for 10 min, then 22 cycles at 95°C for 30 sec, 72°C for 4 min decreasing 1°C each cycle, 68°C for 2 min, followed by 30 cycles at 95°C for 30 sec, 55°C for 30 sec, 68°C for 60 sec. The reaction is loaded onto a 2 % agarose gel. From the tube that gives a PCR product of the correct size, 5  $\mu$ l is used to set up five 1:10 dilutions that are plated onto LB agar + 10 mM  $\text{MgSO}_4$  plates and incubated overnight. A BA85 nitrocellulose filter (Schleicher & Schuell) is placed on top of each plate for 1 hour. The filter is removed, placed phage side up in a petri dish, and covered with 4 ml of SM for 15 min to elute the phage. One milliliter of SM is removed from each plate and used to set up a PCR reaction as above. The plate of the lowest

dilution to give a PCR product is subdivided, filter-lifted and the PCR reaction is repeated. The series of dilutions and subdividing of the plate was continued until a single plaque was isolated that gave a positive PCR band. Once a single plaque was isolated, 10 µl phage supernatant is added to 100 µl SM and 200 µl of K802 cells per plate with a total of 8 plates set up. The plates are incubated overnight at 37°C. The top agarose is removed by adding 8 ml of SM, then scrapping off the agarose with a microscope slide and collecting in a centrifuge tube. To the tube, 3 drops of chloroform is added, vortexed, incubated at 37°C for 15 min then centrifuged for 20 min at 4000 RPM (Sorvall RT6000 table top centrifuge) to recover the phage, which is used to isolate genomic phage DNA using the Qiagen Lambda Midi Kit. The sequences for primers may be derived from the sequences given herein.

For subcloning the coding region of Con-218, PCR is performed in a 50 µl reaction containing 33 µl H<sub>2</sub>O, 5 µl 10X TT buffer (140 mM Ammonium Sulfate, 0.1 % gelatin, 0.6 M Tris-tricine pH 8.4), 5 µl 15 mM MgSO<sub>4</sub>, 2 µl 10 mM dNTP, 4 µl genomic phage DNA (0.1 µg/µl), 0.3 µl LW1643 (1 µg/µl), 0.3 µl LW1644 (1 µg/µl), 0.4 µl High Fidelity Taq polymerase (Boehringer Mannheim). The PCR reaction is started with 1 cycle of 94°C for 2 min followed by 15 cycles at 94°C for 30 sec, 55°C for 60 sec., and 68°C for 2 min. The PCR reaction is loaded onto a 2 % agarose gel. The DNA band is excised from the gel, placed in GenElute Agarose spin column (Supelco) and spun for 10 min at maximum speed. The eluted DNA is EtOH precipitated and resuspended in 12µl H<sub>2</sub>O for ligation. The PCR primer sequences may be derived from the sequences provided herein.

The ligation reaction uses solutions from the TOPO TA Cloning Kit (Invitrogen) which consisted of 4µl PCR product DNA and 1 µl pCRII-TOPO vector that was incubated for 5 minutes at room temperature. To the ligation reaction one microliter of 6X TOPO Cloning Stop Solution is added then the reaction is placed on ice. Two microliters of the ligation reaction is transformed in One-Shot TOP10 cells (Invitrogen), and placed on ice for 30 minutes. The cells are heat-shocked for 30 seconds at 42°C, placed on ice for two minutes, 250 µl of SOC is added,

then incubated at 37°C with shaking for one hour and then plated onto ampicillin plates. A single colony containing an insert is used to inoculate a 5 ml culture of LB medium. Plasmid DNA is purified using a Concert Rapid Plasmid Miniprep System (GibcoBRL) and then sequenced.

Con-218 genomic phage DNA may be sequenced using the ABI PRISM™ 310 Genetic Analyzer (PE Applied Biosystems) which uses advanced capillary electrophoresis technology and the ABI PRISM™ BigDye™ Terminator Cycle Sequencing Ready Reaction Kit. The cycle-sequencing reaction may contain 14 µl of H<sub>2</sub>O, 16 µl of BigDye Terminator mix, 7 µl genomic phage DNA (0.1 µg/µl), and 3 µl primer (25 ng/µl). The reaction is performed in a Perkin-Elmer 9600 thermocycler at 95°C for 5 min, followed by 99 cycles of 95°C for 30 sec, 55°C for 20 sec, and 60°C for 4 min. The product is purified using a Centriflex™ gel filtration cartridges, dried under vacuum, then dissolved in 16 µl of Template Suppression Reagent. The samples are heated at 95°C for 5 min then placed in the 310 Genetic Analyzer

The DNA subcloned into pCRII is sequenced using the ABI PRISM™ 310 Genetic Analyzer (PE Applied Biosystems) which uses advanced capillary electrophoresis technology and the ABI PRISM™ BigDye™ Terminator Cycle Sequencing Ready Reaction Kit. Each cycle-sequencing reaction contains 6 µl of H<sub>2</sub>O, 8 µl of BigDye Terminator mix, 5 µl mini-prep DNA (0.1 µg/µl), and 1 µl primer (25 ng/µl) and is performed in a Perkin-Elmer 9600 thermocycler with 25 cycles of 96°C for 10 sec, 50°C for 10 sec, and 60°C for 4 min. The product is purified using a Centriflex™ gel filtration cartridge, dried under vacuum, then dissolved in 16 µl of Template Suppression Reagent (PE Applied Biosystems). The samples are heated at 95°C for 5 min then placed in the 310 Genetic Analyzer.

#### **Example 5: Hybridization Analysis to demonstrate Con-218 expression in brain**

The expression of Con-218 in mammals, such as the rat, may be investigated by *in situ* hybridization histochemistry. To investigate expression in the brain, for example, coronal and

sagittal rat brain cryosections (20  $\mu$ m thick) are prepared using a Reichert-Jung cryostat. Individual sections are thaw-mounted onto silanized, nuclease-free slides (CEL Associates, Inc., Houston, TX), and stored at -80°C. Sections are processed starting with post-fixation in cold 4% paraformaldehyde, rinsed in cold phosphate-buffered saline (PBS), acetylated using acetic anhydride in triethanolamine buffer, and dehydrated through a series of alcohol washes in 70%, 95%, and 100% alcohol at room temperature. Subsequently, sections are delipidated in chloroform, followed by rehydration through successive exposure to 100% and 95% alcohol at room temperature. Microscope slides containing processed cryosections are allowed to air dry prior to hybridization. Other tissues may be assayed in a similar fashion.

*In situ* hybridization was performed on rat brain sagittal and coronal cryosections, using the following <sup>35</sup>S-labeled oligonucleotides probes designed from the Con-218-RN coding sequence:

Probe 1:

5' -TCGAAAGTCAACAGCAGGCGGTGGCTCCCGCTTAGGGCTGGGTAGGGG-3' [SEQ ID NO. 7]

Probe 2:

5' -GCGGCTACTGAGAGACTCAGAGCGTTGCCGGGGCGTCTCCATCCACGT-3' [SEQ ID NO. 8]

Microscope slides containing sequential brain cryosections were independently exposed to hybridization solution and silanized cover slips were placed over the sections being exposed to hybridization solution. Sections were incubated overnight (15-18 hours) at 52°C to allow hybridization to occur. Equivalent series of cryosections were exposed to antisense Con-218-RN-specific oligonucleotides.

Following the hybridization period, coverslips are washed off the slides in 1X SSC. Following the series of washes, cryosections are dehydrated by consecutive exposure to 70%,

95%, and 100% ammonium acetate in alcohol, followed by air drying and exposure to Kodak BioMax™ MR-1 film. After several weeks of exposure, the film was developed. Data from the above procedure show expression of Con-218-RN in the hypothalamus, cortex piriformis, dentate gyrus, hippocampus and the median habenular nucleus with both oligonucleotide probes.

Expression of Con-218 in the brain provides an indication that modulators of Con-218 activity have utility for treating neurological disorders, including but not limited to, schizophrenia, affective disorders, ADHD/ADD (*i.e.*, Attention Deficit-Hyperactivity Disorder/Attention Deficit Disorder), and neural disorders such as Alzheimer's disease, Parkinson's disease, migraine, and senile dementia. Some other diseases for which modulators of Con-218 may have utility include depression, anxiety, bipolar disease, epilepsy, neuritis, neurasthenia, neuropathy, neuroses, metabolic and cardiovascular diseases and disorders (e.g., type 2 diabetes, obesity, anorexia, dyslipidemias, hypertension, and the like.

Use of Con-218 modulators, including Con-218 ligands and anti-Con-218 antibodies, to treat individuals having such disease states is intended as an aspect of the invention.

#### **Example 6: Tissue Expression Profiling**

Tissue specific expression of the gene encoding Con-218-RN was detected using a PCR-based method. Brain, testis, liver, pancreas and skeletal muscle cDNA was prepared from commercially available rat polyA<sup>+</sup> RNA from Clontech. In the case of rat hypothalamus, polyA<sup>+</sup> RNA was extracted using the Dynabeads mRNA Direct kit (DynaL AS, catalog # 61011). First strand cDNA synthesis was performed according to the description for Superscript Choice System (Life Technologies catalog #18090-019) on 4 microgram polyA<sup>+</sup> RNA using an equal amount of oligo(dT)12-18 and random hexamer primers (Life Technologies catalog # 18418-012 and 48190-011) and Superscript II reverse transcriptase (Life Technologies, catalog # 18064-22). Control PCR templates for ruling out contamination of genomic DNA were generated in the same way, omitting the Superscript II enzyme.

Two Con-218-RN specific oligonucleotide primers were designed and used to amplify a 160 base-pair fragment:

R2A sense primer:

5' - GCCTCTACCTGTTCTGA - 3' [SEQ ID NO. 9]

R2B anti sense primer:

5' - TTTACCCACACGCCACC - 3' [SEQ ID NO. 10]

Each polymerase chain reaction of 50 microliters contained

25 ng cDNA; 200 nM of each primer; 0.2 mM dNTPs (Roche, catalog # 1 969 064); 2.5 U Taq DNA polymerase (Amerham Pharmacia Biotech catalog # 27-0799-62); 50 mM KCl; 1.5 mM MgCl<sub>2</sub>; 10 mM Tris-HCl pH 9.0. Cycling was performed in a MJ Research PTC100 Thermocycler using the following parameters: 94 C for 5 min; 94 C for 1 min; 30 cycles of 55 C for 1 min followed by 72 C for 1 min; 72 C for 5 min.

Analysis was performed on 4% agarose pre-cast gels containing ethidium bromide (E-Gel, Invitrogen #G5018-04). A single band of 160 bp, proven to be correct by subsequent sequencing, was detected in brain, hypothalamus and testis. No amplification products were detected in cDNA from liver, pancreas or skeletal muscle.

A human Multiple Tissue Expression panel (MTE), carrying normalized polyA<sup>+</sup> RNA from 75 different human organs and brain parts, was used for analyzing tissue-specific expression of human Con-218. A 400 base pair fragment was amplified from human genomic DNA using Taq polymerase with the following primers:

H218 sense:

5' - TGGCCCTCTTCTTCATGCTCC - 3' [SEQ ID NO: 11]

and H218 anti-sense:

5' - TTCACCCACACCCACCAG - 3' [SEQ ID NO: 12]

The PCR fragment was excised from a 1.5% agarose gel and purified. The correct identity of the DNA was confirmed by DNA sequencing from both ends using the amplification primers. The probe was randomly labeled with  $\alpha$ -<sup>32</sup>P-dATP (3000 Ci/mmol) using the Strip-Ez Kit, Ambion, Austin, TX, U.S.A, according to the supplied instructions. Non-incorporated nucleotides were removed by spin column purification using Probe Quant G-50 columns (Amersham Pharmacia Biotech, Uppsala, Sweden). The blot was pre-hybridized for 40 minutes at 65°C in 10 ml Express-Hyb solution (Clontech Laboratories, Palo Alto, CA, USA) supplemented with 100µg/ml sonicated salmon sperm DNA (ssDNA). Hybridization was allowed to proceed over-night in fresh solution at 65°C. The filter was washed five times in 2 x SSC/0.1%SDS for 20 minutes at 65°C, and twice in 0.1 x SSC/0.5% SDS for 20 minutes at 55 °C. The membrane was exposed to a phosphoimager-intensifying screen for two days. The results indicated a significant expression in the brain, pituitary, lung, kidney, liver and thyroid gland. In the brain, Con-218 was significantly expressed in the cerebral cortex, cerebellum, amygdala, caudate nucleus, hippocampus, putamen, nucleus accumbens, and thalamus.

#### **Example 7: Northern Blot Analysis**

Northern blots are performed to examine the expression of mRNA. The sense orientation oligonucleotide and the antisense-orientation oligonucleotide, described above, are used as primers to amplify a portion of the GPCR cDNA sequence of a nucleotide sequence of SEQ ID NO: 1.

Multiple human tissue northern blot from Clontech (Human II # 7767-1) are hybridized with the probe. Pre-hybridization is carried out at 42°C for 4 hours in 5xSSC, 1x Denhardt's reagent, 0.1% SDS, 50% formamide, 250 mg/ml salmon sperm DNA. Hybridization is performed overnight at 42°C in the same mixture with the addition of about 1.5x10<sup>6</sup> cpm/ml of labeled probe.



The probe is labeled with  $\alpha$ -<sup>32</sup>P-dCTP by Rediprime™ DNA labeling system (Amersham Pharmacia), purified on Nick Column™ (Amersham Pharmacia) and added to the hybridization solution. The filters are washed several times at 42°C in 0.2x SSC, 0.1% SDS. Filters are exposed to Kodak XAR film (Eastman Kodak Company, Rochester, N.Y., USA) with intensifying screen at -80°C.

### **Example 8: Recombinant Expression of Con-218 in Eukaryotic Host Cells**

#### **A. Expression of Con-218 in Mammalian Cells**

To produce Con-218 protein, a Con-218-encoding polynucleotide is expressed in a suitable host cell using a suitable expression vector and standard genetic engineering techniques. For example, the Con-218-encoding sequence described in Example 1 is subcloned into the commercial expression vector pzeoSV2 (Invitrogen, San Diego, CA) and transfected into Chinese Hamster Ovary (CHO) cells using the transfection reagent FuGENE6™ (Boehringer-Mannheim) and the transfection protocol provided in the product insert. Other eukaryotic cell lines, including human embryonic kidney (HEK 293) and COS cells, are suitable as well. Cells stably expressing Con-218 are selected by growth in the presence of 100 µg/ml zeocin (Stratagene, LaJolla, CA). Optionally, Con-218 may be purified from the cells using standard chromatographic techniques. To facilitate purification, antisera is raised against one or more synthetic peptide sequences that correspond to portions of the Con-218 amino acid sequence, and the antisera is used to affinity purify Con-218. The Con-218 also may be expressed in-frame with a tag sequence (*e.g.*, polyhistidine, hemagglutinin, FLAG) to facilitate purification. Moreover, it will be appreciated that many of the uses for Con-218 polypeptides, such as assays described below, do not require purification of Con-218 from the host cell.

#### **B. Expression of Con-218 in 293 cells**

For expression of Con-218 in mammalian cells 293 (transformed human, primary embryonic kidney cells), a plasmid bearing the relevant Con-218 coding sequence is prepared, using vector pSecTag2A (Invitrogen). Vector pSecTag2A contains the murine IgK chain leader sequence for secretion, the c-myc epitope for detection of the recombinant protein with the anti-myc antibody, a C-terminal polyhistidine for purification with nickel chelate chromatography, and a Zeocin resistant gene for selection of stable transfectants. The forward primer for amplification of this GPCR cDNA is determined by routine procedures and preferably contains a 5' extension of nucleotides to introduce the *HindIII* cloning site and nucleotides matching the GPCR sequence. The reverse primer is also determined by routine procedures and preferably contains a 5' extension of nucleotides to introduce an *XhoI* restriction site for cloning and nucleotides corresponding to the reverse complement of the Con-218 sequence. The PCR conditions are 55°C as the annealing temperature. The PCR product is gel purified and cloned into the *HindIII-XhoI* sites of the vector.

The DNA is purified using Qiagen chromatography columns and transfected into 293 cells using DOTAP™ transfection media (Boehringer Mannheim, Indianapolis, IN). Transiently transfected cells are tested for expression after 24 hours of transfection, using western blots probed with anti-His and anti-Con-218 peptide antibodies. Permanently transfected cells are selected with Zeocin and propagated. Production of the recombinant protein is detected from both cells and media by western blots probed with anti-His, anti-Myc or anti-GPCR peptide antibodies.

### C. Expression of Con-218 in COS cells

For expression of the Con-218 in COS7 cells, a polynucleotide molecule having a nucleotide sequence of SEQ ID NO: 1 can be cloned into vector p3-CI. This vector is a pUC18-derived plasmid that contains the HCMV (human cytomegalovirus) promoter-intron located upstream from the bGH (bovine growth hormone) polyadenylation sequence and a multiple cloning site. In addition, the plasmid contains the dhfr (dihydrofolate reductase) gene which

provides selection in the presence of the drug methotrexane (MTX) for selection of stable transformants.

The forward primer is determined by routine procedures and preferably contains a 5' extension which introduces an *XbaI* restriction site for cloning, followed by nucleotides which correspond to a nucleotide sequence of SEQ ID NO: 1. The reverse primer is also determined by routine procedures and preferably contains 5'- extension of nucleotides which introduces a *SalI* cloning site followed by nucleotides which correspond to the reverse complement of a nucleotide sequence of SEQ ID NO: 1.

The PCR consists of an initial denaturation step of 5 min at 95°C, 30 cycles of 30 sec denaturation at 95°C, 30 sec annealing at 58°C and 30 sec extension at 72°C, followed by 5 min extension at 72°C. The PCR product is gel purified and ligated into the *XbaI* and *SalI* sites of vector p3-CI. This construct is transformed into *E. coli* cells for amplification and DNA purification. The DNA is purified with Qiagen chromatography columns and transfected into COS 7 cells using Lipofectamine™ reagent from BRL, following the manufacturer's protocols. Forty-eight and 72 hours after transfection, the media and the cells are tested for recombinant protein expression.

Con-218 expressed from a COS cell culture can be purified by concentrating the cell-growth media to about 10 mg of protein/ml, and purifying the protein by, for example, chromatography. Purified Con-218 is concentrated to 0.5 mg/ml in an Amicon concentrator fitted with a YM-10 membrane and stored at -80°C.

#### D. Expression of Con-218 in Insect Cells

For expression of Con-218 in a baculovirus system, a polynucleotide molecule having a nucleotide sequence of SEQ ID NO: 1 can be amplified by PCR. The forward primer is determined by routine procedures and preferably contains a 5' extension which adds the *NdeI* cloning site, followed by nucleotides which correspond to a nucleotide sequence of SEQ ID NO:

1. The reverse primer is also determined by routine procedures and preferably contains a 5' extension which introduces the *KpnI* cloning site, followed by nucleotides which correspond to the reverse complement of a nucleotide sequence of SEQ ID NO: 1.

The PCR product is gel purified, digested with *NdeI* and *KpnI*, and cloned into the corresponding sites of vector pACHTL-A (Pharmingen, San Diego, CA). The pACHTL expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV), and a 6XHis tag upstream from the multiple cloning site. A protein kinase site for phosphorylation and a thrombin site for excision of the recombinant protein precede the multiple cloning site is also present. Of course, many other baculovirus vectors could be used in place of pACHTL-A, such as pAc373, pVL941 and pAcIM1. Other suitable vectors for the expression of GPCR polypeptides can be used, provided that the vector construct includes appropriately located signals for transcription, translation, and trafficking, such as an in-frame AUG and a signal peptide, as required. Such vectors are described in Luckow *et al.*, Virology 170:31-39, among others.

The virus is grown and isolated using standard baculovirus expression methods, such as those described in Summers *et al.* (A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures, Texas Agricultural Experimental Station Bulletin No. 1555 (1987)).

In a preferred embodiment, pACHTL-A containing Con-218 gene is introduced into baculovirus using the "BaculoGold™" transfection kit (Pharmingen, San Diego, CA) using methods established by the manufacturer. Individual virus isolates are analyzed for protein production by radiolabeling infected cells with <sup>35</sup>S-methionine at 24 hours post infection. Infected cells are harvested at 48 hours post infection, and the labeled proteins are visualized by SDS-PAGE. Viruses exhibiting high expression levels can be isolated and used for scaled up expression.

For expression of a Con-218 polypeptide in a Sf9 cells, a polynucleotide molecule having the nucleotide sequence of SEQ ID NO: 1 can be amplified by PCR using the primers and

methods described above for baculovirus expression. The Con-218 cDNA is cloned into vector pAcHLT-A (Pharmingen) for expression in Sf9 insect. The insert is cloned into the *NdeI* and *KpnI* sites, after elimination of an internal *NdeI* site (using the same primers described above for expression in baculovirus). DNA is purified with Qiagen chromatography columns and expressed in Sf9 cells. Preliminary Western blot experiments from non-purified plaques are tested for the presence of the recombinant protein of the expected size which reacted with the GPCR-specific antibody. These results are confirmed after further purification and expression optimization in HiG5 cells.

#### Example 9: Interaction Trap/Two-Hybrid System

In order to assay for Con-218-interacting proteins, the interaction trap/two-hybrid library screening method can be used. This assay was first described in Fields *et al.*, *Nature*, **1989**, *340*, 245, which is incorporated herein by reference in its entirety. A protocol is published in Current Protocols in Molecular Biology 1999, John Wiley & Sons, NY and Ausubel, F. M. *et al.* 1992, Short protocols in molecular biology, fourth edition, Greene and Wiley-interscience, NY, which is incorporated herein by reference in its entirety. Kits are available from Clontech, Palo Alto, CA (Matchmaker Two-Hybrid System 3).

A fusion of the nucleotide sequences encoding all or partial Con-218 and the yeast transcription factor GAL4 DNA-binding domain (DNA-BD) is constructed in an appropriate plasmid (ie. pGBKT7) using standard subcloning techniques. Similarly, a GAL4 active domain (AD) fusion library is constructed in a second plasmid (i.e., pGADT7) from cDNA of potential GPCR-binding proteins (for protocols on forming cDNA libraries, see Sambrook et al. 1989, Molecular cloning: a laboratory manual, second edition, Cold Spring Harbor Press, Cold Spring Harbor, NY), which is incorporated herein by reference in its entirety. The DNA-BD/GPCR fusion construct is verified by sequencing, and tested for autonomous reporter gene activation and cell toxicity, both of which would prevent a successful two-hybrid analysis. Similar controls are performed with the AD/library fusion construct to ensure expression in host cells and lack of

transcriptional activity. Yeast cells are transformed (ca. 10<sup>5</sup> transformants/mg DNA) with both the GPCR and library fusion plasmids according to standard procedure (Ausubel, et al., 1992, Short protocols in molecular biology, fourth edition, Greene and Wiley-interscience, NY, which is incorporated herein by reference in its entirety). In vivo binding of DNA-BD/GPCR with AD/library proteins results in transcription of specific yeast plasmid reporter genes (i.e., lacZ, HIS3, ADE2, LEU2). Yeast cells are plated on nutrient-deficient media to screen for expression of reporter genes. Colonies are dually assayed for  $\beta$ -galactosidase activity upon growth in Xgal (5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactoside) supplemented media (filter assay for  $\beta$ -galactosidase activity is described in Breeden *et al.*, Cold Spring Harb. Symp. Quant. Biol., 1985, 50, 643, which is incorporated herein by reference in its entirety). Positive AD-library plasmids are rescued from transformants and reintroduced into the original yeast strain as well as other strains containing unrelated DNA-BD fusion proteins to confirm specific Con-218/library protein interactions. Insert DNA is sequenced to verify the presence of an open reading frame fused to GAL4 AD and to determine the identity of the Con-218-binding protein.

#### Example 10: Antibodies to Con-218

Standard techniques are employed to generate polyclonal or monoclonal antibodies to the Con-218 receptor, and to generate useful antigen-binding fragments thereof or variants thereof, including "humanized" variants. Such protocols can be found, for example, in Sambrook et al. (1989) and Harlow *et al.* (Eds.), Antibodies A Laboratory Manual; Cold Spring Harbor Laboratory; Cold Spring Harbor, NY (1988). In one embodiment, recombinant Con-218 polypeptides (or cells or cell membranes containing such polypeptides) are used as antigen to generate the antibodies. In another embodiment, one or more peptides having amino acid sequences corresponding to an immunogenic portion of Con-218 (e.g., 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more amino acids) are used as antigen. Peptides corresponding to

extracellular portions of Con-218, especially hydrophilic extracellular portions, are preferred. The antigen may be mixed with an adjuvant or linked to a hapten to increase antibody production.

A. Polyclonal or Monoclonal antibodies

As one exemplary protocol, recombinant Con-218 or a synthetic fragment thereof is used to immunize a mouse for generation of monoclonal antibodies (or larger mammal, such as a rabbit, for polyclonal antibodies). To increase antigenicity, peptides are conjugated to Keyhole Limpet Hemocyanin (Pierce), according to the manufacturer's recommendations. For an initial injection, the antigen is emulsified with Freund's Complete Adjuvant and injected subcutaneously. At intervals of two to three weeks, additional aliquots of Con-218 antigen are emulsified with Freund's Incomplete Adjuvant and injected subcutaneously. Prior to the final booster injection, a serum sample is taken from the immunized mice and assayed by western blot to confirm the presence of antibodies that immunoreact with Con-218. Serum from the immunized animals may be used as polyclonal antisera or used to isolate polyclonal antibodies that recognize Con-218. Alternatively, the mice are sacrificed and their spleen removed for generation of monoclonal antibodies.

To generate monoclonal antibodies, the spleens are placed in 10 ml serum-free RPMI 1640, and single cell suspensions are formed by grinding the spleens in serum-free RPMI 1640, supplemented with 2 mM L-glutamine, 1 mM sodium pyruvate, 100 units/ml penicillin, and 100 µg/ml streptomycin (RPMI) (Gibco, Canada). The cell suspensions are filtered and washed by centrifugation and resuspended in serum-free RPMI. Thymocytes taken from three naive Balb/c mice are prepared in a similar manner and used as a Feeder Layer. NS-1 myeloma cells, kept in log phase in RPMI with 10% fetal bovine serum (FBS) (Hyclone Laboratories, Inc., Logan, Utah) for three days prior to fusion, are centrifuged and washed as well.

To produce hybridoma fusions, spleen cells from the immunized mice are combined with NS-1 cells and centrifuged, and the supernatant is aspirated. The cell pellet is dislodged by

tapping the tube, and 2 ml of 37°C PEG 1500 (50% in 75 mM HEPES, pH 8.0) (Boehringer-Mannheim) is stirred into the pellet, followed by the addition of serum-free RPMI. Thereafter, the cells are centrifuged, resuspended in RPMI containing 15% FBS, 100  $\mu$ M sodium hypoxanthine, 0.4  $\mu$ M aminopterin, 16  $\mu$ M thymidine (HAT) (Gibco), 25 units/ml IL-6 (Boehringer-Mannheim) and  $1.5 \times 10^6$  thymocytes/ml, and plated into 10 Corning flat-bottom 96-well tissue culture plates (Corning, Corning New York).

On days 2, 4, and 6 after the fusion, 100  $\mu$ l of medium is removed from the wells of the fusion plates and replaced with fresh medium. On day 8, the fusions are screened by ELISA, testing for the presence of mouse IgG that binds to Con-218. Selected fusion wells are further cloned by dilution until monoclonal cultures producing anti-Con-218 antibodies are obtained.

#### B. Humanization of anti-Con-218 monoclonal antibodies

The expression pattern of Con-218 as reported herein and the proven track record of GPCRs as targets for therapeutic intervention suggest therapeutic indications for Con-218 inhibitors (antagonists). Con-218-neutralizing antibodies comprise one class of therapeutics useful as Con-218 antagonists. Following are protocols to improve the utility of anti-Con-218 monoclonal antibodies as therapeutics in humans by "humanizing" the monoclonal antibodies to improve their serum half-life and render them less immunogenic in human hosts (*i.e.*, to prevent human antibody response to non-human anti-Con-218 antibodies).

The principles of humanization have been described in the literature and are facilitated by the modular arrangement of antibody proteins. To minimize the possibility of binding complement, a humanized antibody of the IgG4 isotype is preferred.

For example, a level of humanization is achieved by generating chimeric antibodies comprising the variable domains of non-human antibody proteins of interest with the constant domains of human antibody molecules. (See, *e.g.*, Morrison *et al.*, Adv. Immunol., 44:65-92 (1989)). The variable domains of Con-218-neutralizing anti-Con-218 antibodies are cloned from



the genomic DNA of a B-cell hybridoma or from cDNA generated from mRNA isolated from the hybridoma of interest. The V region gene fragments are linked to exons encoding human antibody constant domains, and the resultant construct is expressed in suitable mammalian host cells (*e.g.*, myeloma or CHO cells).

To achieve an even greater level of humanization, only those portions of the variable region gene fragments that encode antigen-binding complementarity determining regions ("CDR") of the non-human monoclonal antibody genes are cloned into human antibody sequences. (See, *e.g.*, Jones *et al.*, Nature 321:522-525 (1986); Riechmann *et al.*, Nature 332:323-327 (1988); Verhoeyen *et al.*, Science 239:1534-36 (1988); and Tempest *et al.*, Bio/Technology 9:266-71 (1991)). If necessary, the  $\beta$ -sheet framework of the human antibody surrounding the CDR3 regions also is modified to more closely mirror the three dimensional structure of the antigen-binding domain of the original monoclonal antibody. (See Kettleborough *et al.*, Protein Engin., 4:773-783 (1991); and Foote *et al.*, J. Mol. Biol., 224:487-499 (1992)).

In an alternative approach, the surface of a non-human monoclonal antibody of interest is humanized by altering selected surface residues of the non-human antibody, *e.g.*, by site-directed mutagenesis, while retaining all of the interior and contacting residues of the non-human antibody. See Padlan, Molecular Immunol., 28(4/5):489-98 (1991).

The foregoing approaches are employed using Con-218-neutralizing anti-Con-218 monoclonal antibodies and the hybridomas that produce them to generate humanized Con-218-neutralizing antibodies useful as therapeutics to treat or palliate conditions wherein Con-218 expression or ligand-mediated Con-218 signaling is detrimental.

### C. Human Con-218-Neutralizing Antibodies from Phage Display

Human Con-218-neutralizing antibodies are generated by phage display techniques such as those described in Aujame *et al.*, Human Antibodies 8(4):155-168 (1997); Hoogenboom,

TIBTECH 15:62-70 (1997); and Rader *et al.*, Curr. Opin. Biotechnol. 8:503-508 (1997), all of which are incorporated by reference. For example, antibody variable regions in the form of Fab fragments or linked single chain Fv fragments are fused to the amino terminus of filamentous phage minor coat protein pIII. Expression of the fusion protein and incorporation thereof into the mature phage coat results in phage particles that present an antibody on their surface and contain the genetic material encoding the antibody. A phage library comprising such constructs is expressed in bacteria, and the library is screened for Con-218-specific phage-antibodies using labeled or immobilized Con-218 as antigen-probe.

D. Human Con-218-neutralizing antibodies from transgenic mice

Human Con-218-neutralizing antibodies are generated in transgenic mice essentially as described in Bruggemann *et al.*, Immunol. Today 17(8):391-97 (1996) and Bruggemann *et al.*, Curr. Opin. Biotechnol. 8:455-58 (1997). Transgenic mice carrying human V-gene segments in germline configuration and that express these transgenes in their lymphoid tissue are immunized with a Con-218 composition using conventional immunization protocols. Hybridomas are generated using B cells from the immunized mice using conventional protocols and screened to identify hybridomas secreting anti-Con-218 human antibodies (*e.g.*, as described above).

**Example 11: Assays to Identify Modulators of Con-218 Activity**

Set forth below are several nonlimiting assays for identifying modulators (agonists and antagonists) of Con-218 activity. Among the modulators that can be identified by these assays are natural ligand compounds of the receptor; synthetic analogs and derivatives of natural ligands; antibodies, antibody fragments, and/or antibody-like compounds derived from natural antibodies or from antibody-like combinatorial libraries; and/or synthetic compounds identified by high-throughput screening of libraries; and the like. All modulators that bind Con-218 are useful for identifying Con-218 in tissue samples (*e.g.*, for diagnostic purposes, pathological purposes, and the like). Agonist and antagonist modulators are useful for up-regulating and

down-regulating Con-218 activity, respectively, to treat disease states characterized by abnormal levels of Con-218 activity. The assays may be performed using single putative modulators, and/or may be performed using a known agonist in combination with candidate antagonists (or visa versa).

#### A. cAMP Assays

In one type of assay, levels of cyclic adenosine monophosphate (cAMP) are measured in Con-218-transfected cells that have been exposed to candidate modulator compounds. Protocols for cAMP assays have been described in the literature. (See, *e.g.*, Sutherland *et al.*, *Circulation* 37: 279 (1968); Frandsen *et al.*, *Life Sciences* 18: 529-541 (1976); Dooley *et al.*, *Journal of Pharmacology and Experimental Therapeutics* 283 (2): 735-41 (1997); and George *et al.*, *Journal of Biomolecular Screening* 2 (4): 235-40 (1997)). An exemplary protocol for such an assay, using an Adenylyl Cyclase Activation FlashPlate® Assay from NEN™ Life Science Products, is set forth below.

Briefly, the Con-218 coding sequence (*e.g.*, a cDNA or intronless genomic DNA) is subcloned into a commercial expression vector, such as pzeoSV2 (Invitrogen), and transiently transfected into Chinese Hamster Ovary (CHO) cells using known methods, such as the transfection protocol provided by Boehringer-Mannheim when supplying the FuGENE 6 transfection reagent. Transfected CHO cells are seeded into 96-well microplates from the FlashPlate® assay kit, which are coated with solid scintillant to which antisera to cAMP has been bound. For a control, some wells are seeded with wild type (untransfected) CHO cells. Other wells in the plate receive various amounts of a cAMP standard solution for use in creating a standard curve.

One or more test compounds (*i.e.*, candidate modulators) are added to the cells in each well, with water and/or compound-free medium/diluent serving as a control or controls. After treatment, cAMP is allowed to accumulate in the cells for exactly 15 minutes at room temperature. The assay is terminated by the addition of lysis buffer containing [<sup>125</sup>I]-labeled

cAMP, and the plate is counted using a Packard Topcount™ 96-well microplate scintillation counter. Unlabeled cAMP from the lysed cells (or from standards) and fixed amounts of [<sup>125</sup>I]-cAMP compete for antibody bound to the plate. A standard curve is constructed, and cAMP values for the unknowns are obtained by interpolation. Changes in intracellular cAMP levels of cells in response to exposure to a test compound are indicative of Con-218 modulating activity. Modulators that act as agonists of receptors which couple to the G<sub>s</sub> subtype of G proteins will stimulate production of cAMP, leading to a measurable 3-10 fold increase in cAMP levels. Agonists of receptors which couple to the G<sub>i/o</sub> subtype of G proteins will inhibit forskolin-stimulated cAMP production, leading to a measurable decrease in cAMP levels of 50-100%. Modulators that act as inverse agonists will reverse these effects at receptors that are either constitutively active or activated by known agonists.

#### B. Aequorin Assays

In another assay, cells (*e.g.*, CHO cells) are transiently co-transfected with both a Con-218 expression construct and a construct that encodes the photoprotein apoaequorin. In the presence of the cofactor coelenterazine, apoaequorin will emit a measurable luminescence that is proportional to the amount of intracellular (cytoplasmic) free calcium. (See generally, Cobbold, et al. "Aequorin measurements of cytoplasmic free calcium," *In*: McCormack J.G. and Cobbold P.H., eds., *Cellular Calcium: A Practical Approach*. Oxford:IRL Press (1991); Stables *et al.*, *Analytical Biochemistry* 252: 115-26 (1997); and Haugland, *Handbook of Fluorescent Probes and Research Chemicals*. Sixth edition. Eugene OR: Molecular Probes (1996).)

In one exemplary assay, Con-218 is subcloned into the commercial expression vector pzeoSV2 (Invitrogen) and transiently co-transfected along with a construct that encodes the photoprotein apoaequorin (Molecular Probes, Eugene, OR) into CHO cells using the transfection reagent FuGENE 6 (Boehringer-Mannheim) and the transfection protocol provided in the product insert.

The cells are cultured for 24 hours at 37°C in MEM (Gibco/BRL, Gaithersburg, MD) supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 µg/ml streptomycin, at which time the medium is changed to serum-free MEM containing 5 µM coelenterazine (Molecular Probes, Eugene, OR). Culturing is then continued for two additional hours at 37°C. Subsequently, cells are detached from the plate using VERSEN (Gibco/BRL), washed, and resuspended at 200,000 cells/ml in serum-free MEM.

Dilutions of candidate Con-218 modulator compounds are prepared in serum-free MEM and dispensed into wells of an opaque 96-well assay plate at 50 µl/well. Plates are then loaded onto an MLX microtiter plate luminometer (Dynex Technologies, Inc., Chantilly, VA). The instrument is programmed to dispense 50 µl cell suspensions into each well, one well at a time, and immediately read luminescence for 15 seconds. Dose-response curves for the candidate modulators are constructed using the area under the curve for each light signal peak. Data are analyzed with SlideWrite, using the equation for a one-site ligand, and EC<sub>50</sub> values are obtained. Changes in luminescence caused by the compounds are considered indicative of modulatory activity. Modulators that act as agonists at receptors which couple to the G<sub>q</sub> subtype of G proteins give an increase in luminescence of up to 100 fold. Modulators that act as inverse agonists will reverse this effect at receptors that are either constitutively active or activated by known agonists.

#### C. Luciferase Reporter Gene Assay

The photoprotein luciferase provides another useful tool for assaying for modulators of Con-218 activity. Cells (*e.g.*, CHO cells or COS 7 cells) are transiently co-transfected with both a Con-218 expression construct (*e.g.*, Con-218 in pzeoSV2) and a reporter construct which includes a gene for the luciferase protein downstream from a transcription factor binding site, such as the cAMP-response element (CRE), AP-1, or NF-kappa B. Agonist binding to receptors coupled to the G<sub>s</sub> subtype of G proteins leads to increases in cAMP, thereby activating the CRE

transcription factor and resulting in expression of the luciferase gene. Agonist binding to receptors coupled to the G<sub>q</sub> subtype of G protein leads to production of diacylglycerol that activates protein kinase C, which activates the AP-1 or NF-kappa B transcription factors, in turn resulting in expression of the luciferase gene. Expression levels of luciferase reflect the activation status of the signaling events. (See generally, George *et al.*, Journal of Biomolecular Screening 2(4): 235-240 (1997); and Stratowa *et al.*, Current Opinion in Biotechnology 6: 574-581 (1995)). Luciferase activity may be quantitatively measured using, *e.g.*, luciferase assay reagents that are commercially available from Promega (Madison, WI).

In one exemplary assay, CHO cells are plated in 24-well culture dishes at a density of 100,000 cells/well one day prior to transfection and cultured at 37°C in MEM (Gibco/BRL) supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 µg/ml streptomycin. Cells are transiently co-transfected with both a Con-218 expression construct and a reporter construct containing the luciferase gene. The reporter plasmids CRE-luciferase, AP-1-luciferase and NF-kappaB-luciferase may be purchased from Stratagene (LaJolla, CA). Transfections are performed using the FuGENE 6 transfection reagent (Boehringer-Mannheim) according to the supplier's instructions. Cells transfected with the reporter construct alone are used as a control. Twenty-four hours after transfection, cells are washed once with PBS pre-warmed to 37°C. Serum-free MEM is then added to the cells either alone (control) or with one or more candidate modulators and the cells are incubated at 37°C for five hours. Thereafter, cells are washed once with ice-cold PBS and lysed by the addition of 100 µl of lysis buffer per well from the luciferase assay kit supplied by Promega. After incubation for 15 minutes at room temperature, 15 µl of the lysate is mixed with 50 µl of substrate solution (Promega) in an opaque-white, 96-well plate, and the luminescence is read immediately on a Wallace model 1450 MicroBeta scintillation and luminescence counter (Wallace Instruments, Gaithersburg, MD).

Differences in luminescence in the presence versus the absence of a candidate modulator compound are indicative of modulatory activity. Receptors that are either constitutively active or

activated by agonists typically give a 3-20-fold stimulation of luminescence compared to cells transfected with the reporter gene alone. Modulators that act as inverse agonists will reverse this effect.

D. Intracellular calcium measurement using FLIPR

Changes in intracellular calcium levels are another recognized indicator of G protein-coupled receptor activity, and such assays can be employed to screen for modulators of Con-218 activity. For example, CHO cells stably transfected with a Con-218 expression vector are plated at a density of  $4 \times 10^4$  cells/well in Packard black-walled, 96-well plates specially designed to discriminate fluorescence signals emanating from the various wells on the plate. The cells are incubated for 60 minutes at 37°C in modified Dulbecco's PBS (D-PBS) containing 36 mg/L pyruvate and 1 g/L glucose with the addition of 1% fetal bovine serum and one of four calcium indicator dyes (Fluo-3™ AM, Fluo-4™ AM, Calcium Green™-1 AM, or Oregon Green™ 488 BAPTA-1 AM), each at a concentration of 4 μM. Plates are washed once with modified D-PBS without 1% fetal bovine serum and incubated for 10 minutes at 37°C to remove residual dye from the cellular membrane. In addition, a series of washes with modified D-PBS without 1% fetal bovine serum is performed immediately prior to activation of the calcium response.

A calcium response is initiated by the addition of one or more candidate receptor agonist compounds, calcium ionophore A23187 (10 μM; positive control), or ATP (4 μM; positive control). Fluorescence is measured by Molecular Device's FLIPR with an argon laser (excitation at 488 nm). (See, *e.g.*, Kuntzweiler *et al.*, Drug Development Research, 44(1):14-20 (1998)). The F-stop for the detector camera was set at 2.5 and the length of exposure was 0.4 milliseconds. Basal fluorescence of cells was measured for 20 seconds prior to addition of candidate agonist, ATP, or A23187, and the basal fluorescence level was subtracted from the response signal. The calcium signal is measured for approximately 200 seconds, taking readings every two seconds. Calcium ionophore A23187 and ATP increase the calcium signal 200%

above baseline levels. In general, activated GPCRs increase the calcium signal approximately 10-15% above baseline signal.

E. Mitogenesis Assay

In a mitogenesis assay, the ability of candidate modulators to induce or inhibit Con-218-mediated cell division is determined. (See, e.g., Lajiness *et al.*, Journal of Pharmacology and Experimental Therapeutics 267(3):1573-1581 (1993)). For example, CHO cells stably expressing Con-218 are seeded into 96-well plates at a density of 5000 cells/well and grown at 37°C in MEM with 10% fetal calf serum for 48 hours, at which time the cells are rinsed twice with serum-free MEM. After rinsing, 80 µl of fresh MEM, or MEM containing a known mitogen, is added along with 20 µl MEM containing varying concentrations of one or more candidate modulators or test compounds diluted in serum-free medium. As controls, some wells on each plate receive serum-free medium alone, and some receive medium containing 10% fetal bovine serum. Untransfected cells or cells transfected with vector alone also may serve as controls.

After culture for 16-18 hours, 1µCi of [<sup>3</sup>H]-thymidine (2 Ci/mmol) is added to the wells and cells are incubated for an additional 2 hours at 37°C. The cells are trypsinized and collected on filter mats with a cell harvester (Tomtec); the filters are then counted in a Betaplate counter. The incorporation of [<sup>3</sup>H]-thymidine in serum-free test wells is compared to the results achieved in cells stimulated with serum (positive control). Use of multiple concentrations of test compounds permits creation and analysis of dose-response curves using the non-linear, least squares fit equation:  $A = B \times [C / (D + C)] + G$  where A is the percent of serum stimulation; B is the maximal effect minus baseline; C is the EC<sub>50</sub>; D is the concentration of the compound; and G is the maximal effect. Parameters B, C and G are determined by Simplex optimization.



Agonists that bind to the receptor are expected to increase [ $^3\text{H}$ ]-thymidine incorporation into cells, showing up to 80% of the response to serum. Antagonists that bind to the receptor will inhibit the stimulation seen with a known agonist by up to 100%.

F. [ $^{35}\text{S}$ ]GTP $\gamma\text{S}$  Binding Assay

Because G protein-coupled receptors signal through intracellular G proteins whose activity involves GTP binding and hydrolysis to yield bound GDP, measurement of binding of the non-hydrolyzable GTP analog [ $^{35}\text{S}$ ]GTP $\gamma\text{S}$  in the presence and absence of candidate modulators provides another assay for modulator activity. (See, e.g., Kowal *et al.*, *Neuropharmacology* 37:179-187 (1998).)

In one exemplary assay, cells stably transfected with a Con-218 expression vector are grown in 10 cm tissue culture dishes to subconfluence, rinsed once with 5 ml of ice-cold  $\text{Ca}^{2+}/\text{Mg}^{2+}$ -free phosphate-buffered saline, and scraped into 5 ml of the same buffer. Cells are pelleted by centrifugation (500 x g, 5 minutes), resuspended in TEE buffer (25 mM Tris, pH 7.5, 5 mM EDTA, 5 mM EGTA), and frozen in liquid nitrogen. After thawing, the cells are homogenized using a Dounce homogenizer (one ml TEE per plate of cells), and centrifuged at 1,000 x g for 5 minutes to remove nuclei and unbroken cells.

The homogenate supernatant is centrifuged at 20,000 x g for 20 minutes to isolate the membrane fraction, and the membrane pellet is washed once with TEE and resuspended in binding buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 10 mM  $\text{MgCl}_2$ , 1 mM EDTA). The resuspended membranes can be frozen in liquid nitrogen and stored at  $-70^\circ\text{C}$  until use.

Aliquots of cell membranes prepared as described above and stored at  $-70^\circ\text{C}$  are thawed, homogenized, and diluted into buffer containing 20 mM HEPES, 10 mM  $\text{MgCl}_2$ , 1 mM EDTA, 120 mM NaCl, 10  $\mu\text{M}$  GDP, and 0.2 mM ascorbate, at a concentration of 10-50  $\mu\text{g}/\text{ml}$ . In a final volume of 90  $\mu\text{l}$ , homogenates are incubated with varying concentrations of candidate modulator compounds or 100  $\mu\text{M}$  GTP for 30 minutes at  $30^\circ\text{C}$  and then placed on ice. To each sample, 10

μl guanosine 5'-O-(3[<sup>35</sup>S]thio) triphosphate (NEN, 1200 Ci/mmol; [<sup>35</sup>S]-GTPγS), was added to a final concentration of 100-200 pM. Samples are incubated at 30°C for an additional 30 minutes, 1 ml of 10 mM HEPES, pH 7.4, 10 mM MgCl<sub>2</sub>, at 4°C is added and the reaction is stopped by filtration.

Samples are filtered over Whatman GF/B filters and the filters are washed with 20 ml ice-cold 10 mM HEPES, pH 7.4, 10 mM MgCl<sub>2</sub>. Filters are counted by liquid scintillation spectroscopy. Nonspecific binding of [<sup>35</sup>S]-GTPγS is measured in the presence of 100 μM GTP and subtracted from the total. Compounds are selected that modulate the amount of [<sup>35</sup>S]-GTPγS binding in the cells, compared to untransfected control cells. Activation of receptors by agonists gives up to a five-fold increase in [<sup>35</sup>S]GTPγS binding. This response is blocked by antagonists.

#### G. MAP Kinase Activity Assay

Evaluation of MAP kinase activity in cells expressing a GPCR provides another assay to identify modulators of GPCR activity. (See, *e.g.*, Lajiness *et al.*, Journal of Pharmacology and Experimental Therapeutics 267(3):1573-1581 (1993) and Boulton *et al.*, Cell 65:663-675 (1991).)

In one embodiment, CHO cells stably transfected with Con-218 are seeded into 6-well plates at a density of 70,000 cells/well 48 hours prior to the assay. During this 48-hour period, the cells are cultured at 37°C in MEM medium supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 μg/ml streptomycin. The cells are serum-starved for 1-2 hours prior to the addition of stimulants.

For the assay, the cells are treated with medium alone or medium containing either a candidate agonist or 200 nM Phorbol ester- myristoyl acetate (*i.e.*, PMA, a positive control), and the cells are incubated at 37°C for varying times. To stop the reaction, the plates are placed on ice, the medium is aspirated, and the cells are rinsed with 1 ml of ice-cold PBS containing 1 mM EDTA. Thereafter, 200 μl of cell lysis buffer (12.5 mM MOPS, pH 7.3, 12.5 mM

glycerophosphate, 7.5 mM MgCl<sub>2</sub>, 0.5 mM EGTA, 0.5 mM sodium vanadate, 1 mM benzamidine, 1 mM dithiothreitol, 10 µg/ml leupeptin, 10 µg/ml aprotinin, 2 µg/ml pepstatin A, and 1 µM okadaic acid) is added to the cells. The cells are scraped from the plates and homogenized by 10 passages through a 23 3/4 G needle, and the cytosol fraction is prepared by centrifugation at 20,000 x g for 15 minutes.

Aliquots (5-10 µl containing 1-5 µg protein) of cytosol are mixed with 1 mM MAPK Substrate Peptide (APRTPGGRR (SEQ ID NO: 7), Upstate Biotechnology, Inc., N.Y.) and 50 µM [ $\gamma$ -<sup>32</sup>P]ATP (NEN, 3000 Ci/mmol), diluted to a final specific activity of ~2000 cpm/pmol, in a total volume of 25 µl. The samples are incubated for 5 minutes at 30°C, and reactions are stopped by spotting 20 µl on 2 cm<sup>2</sup> squares of Whatman P81 phosphocellulose paper. The filter squares are washed in 4 changes of 1% H<sub>3</sub>PO<sub>4</sub>, and the squares are subjected to liquid scintillation spectroscopy to quantitate bound label. Equivalent cytosolic extracts are incubated without MAPK substrate peptide, and the bound label from these samples are subtracted from the matched samples with the substrate peptide. The cytosolic extract from each well is used as a separate point. Protein concentrations are determined by a dye binding protein assay (Bio-Rad Laboratories). Agonist activation of the receptor is expected to result in up to a five-fold increase in MAPK enzyme activity. This increase is blocked by antagonists.

#### H. [<sup>3</sup>H]Arachidonic Acid Release

The activation of GPCRs also has been observed to potentiate arachidonic acid release in cells, providing yet another useful assay for modulators of GPCR activity. (*See, e.g., Kanterman et al., Molecular Pharmacology 39:364-369 (1991).*) For example, CHO cells that are stably transfected with a Con-218 expression vector are plated in 24-well plates at a density of 15,000 cells/well and grown in MEM medium supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 µg/ml streptomycin for 48 hours at 37°C before use. Cells of each well are labeled by incubation with [<sup>3</sup>H]-arachidonic acid (Amersham Corp., 210

Ci/mmol) at 0.5  $\mu$ Ci/ml in 1 ml MEM supplemented with 10 mM HEPES, pH 7.5, and 0.5% fatty-acid-free bovine serum albumin for 2 hours at 37°C. The cells are then washed twice with 1 ml of the same buffer.

Candidate modulator compounds are added in 1 ml of the same buffer, either alone or with 10  $\mu$ M ATP and the cells are incubated at 37°C for 30 minutes. Buffer alone and mock-transfected cells are used as controls. Samples (0.5 ml) from each well are counted by liquid scintillation spectroscopy. Agonists which activate the receptor will lead to potentiation of the ATP-stimulated release of [ $^3$ H]-arachidonic acid. This potentiation is blocked by antagonists.

#### I. Extracellular Acidification Rate

In yet another assay, the effects of candidate modulators of Con-218 activity are assayed by monitoring extracellular changes in pH induced by the test compounds. (*See, e.g., Dunlop et al., Journal of Pharmacological and Toxicological Methods* 40(1):47-55 (1998).) In one embodiment, CHO cells transfected with a Con-218 expression vector are seeded into 12 mm capsule cups (Molecular Devices Corp.) at  $4 \times 10^5$  cells/cup in MEM supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 10 U/ml penicillin, and 10  $\mu$ g/ml streptomycin. The cells are incubated in this medium at 37°C in 5% CO<sub>2</sub> for 24 hours.

Extracellular acidification rates are measured using a Cytosensor microphysiometer (Molecular Devices Corp.). The capsule cups are loaded into the sensor chambers of the microphysiometer and the chambers are perfused with running buffer (bicarbonate-free MEM supplemented with 4 mM L-glutamine, 10 units/ml penicillin, 10  $\mu$ g/ml streptomycin, 26 mM NaCl) at a flow rate of 100  $\mu$ l/minute. Candidate agonists or other agents are diluted into the running buffer and perfused through a second fluid path. During each 60-second pump cycle, the pump is run for 38 seconds and is off for the remaining 22 seconds. The pH of the running buffer in the sensor chamber is recorded during the cycle from 43-58 seconds, and the pump is re-started at 60 seconds to start the next cycle. The rate of acidification of the running buffer

during the recording time is calculated by the Cytosoft program. Changes in the rate of acidification are calculated by subtracting the baseline value (the average of 4 rate measurements immediately before addition of a modulator candidate) from the highest rate measurement obtained after addition of a modulator candidate. The selected instrument detects 61 mV/pH unit. Modulators that act as agonists of the receptor result in an increase in the rate of extracellular acidification compared to the rate in the absence of agonist. This response is blocked by modulators which act as antagonists of the receptor.

#### **Example 12 - Using Con-218 proteins to isolate neurotransmitters**

The isolated Con-218 proteins can be used to isolate novel or known neurotransmitters (Saito *et al.*, Nature 400: 265-269, 1999). The cDNAs that encode the isolated Con-218 can be cloned into mammalian expression vectors and used to stably or transiently transfect mammalian cells including CHO, Cos or HEK293 cells. Receptor expression can be determined by Northern blot analysis of transfected cells and identification of an appropriately sized mRNA band (predicted size from the cDNA). Brain regions shown by mRNA analysis to express each of the Con-218 proteins could be processed for peptide extraction using any of several protocols ((Reinsheidk R.K. *et al.*, Science 270: 243-247, 1996; Sakurai, T., *et al.*, Cell 92; 573-585, 1998; Hinuma, S., *et al.*, Nature 393: 272-276, 1998). Chromatographic fractions of brain extracts could be tested for ability to activate Con-218 proteins by measuring second messenger production such as changes in cAMP production in the presence or absence of forskolin, changes in inositol 3-phosphate levels, changes in intracellular calcium levels or by indirect measures of receptor activation including receptor stimulated mitogenesis, receptor mediated changes in extracellular acidification or receptor mediated changes in reporter gene activation in response to cAMP or calcium (these methods should all be referenced in other sections of the patent). Receptor activation could also be monitored by co-transfecting cells with a chimeric  $GI_{q/13}$  to force receptor coupling to a calcium stimulating pathway (Conklin *et al.*, Nature 363; 274-276,

1993). Neurotransmitter mediated activation of receptors could also be monitored by measuring changes in [<sup>35</sup>S]-GTPKS binding in membrane fractions prepared from transfected mammalian cells. This assay could also be performed using baculoviruses containing Con-218 proteins infected into SF9 insect cells.

The neurotransmitter which activates Con-218 proteins can be purified to homogeneity through successive rounds of purification using Con-218 protein activation as a measurement of neurotransmitter activity. The composition of the neurotransmitter can be determined by mass spectrometry and Edman degradation if peptidergic. Neurotransmitters isolated in this manner will be bioactive materials which will alter neurotransmission in the central nervous system and will produce behavioral and biochemical changes.

#### **Example 13 - Using Con-218 proteins to isolate and purify G proteins**

cDNAs encoding Con-218 proteins are epitope-tagged at the amino terminus end of the cDNA with the cleavable influenza-hemagglutinin signal sequence followed by the FLAG epitope (IBI, New Haven, CT). Additionally, these sequences are tagged at the carboxyl terminus with DNA encoding six histidine residues. (Amino and Carboxyl Terminal Modifications to Facilitate the Production and Purification of a G Protein-Coupled Receptor, B.K. Kobilka, *Analytical Biochemistry*, Vol. 231, No. 1, Oct 1995, pp. 269-271). The resulting sequences are cloned into a baculovirus expression vector such as pVL1392 (Invitrogen). The baculovirus expression vectors are used to infect SF-9 insect cells as described (Guan, X. M., Kobilka, T. S., and Kobilka, B. K. (1992) *J. Biol. Chem.* **267**, 21995-21998). Infected SF-9 cells could be grown in 1000-ml cultures in SF900 II medium (Life Technologies, Inc.) containing 5% fetal calf serum (Gemini, Calabasas, CA) and 0.1 mg/ml gentamicin (Life Technologies, Inc.) for 48 hours at which time the cells could be harvested. Cell membrane preparations could be separated from soluble proteins following cell lysis. Con-218 protein purification is carried out as described for purification of the 92 receptor (Kobilka, *Anal. Biochem.*, 231 (1): 269-271,

1995) including solubilization of the membranes in 0.8-1.0 % *n*-dodecyl -D-maltoside (DM) (CalBiochem, La Jolla, CA) in buffer containing protease inhibitors followed by Ni-column chromatography using chelating Sepharose™ (Pharmacia, Uppsala, Sweden). The eluate from the Ni-column is further purified on an M1 anti-FLAG antibody column (IBI). Receptor containing fractions are monitored by using receptor specific antibodies following western blot analysis or by SDS-PAGE analysis to look for an appropriate sized protein band (appropriate size would be the predicted molecular weight of the protein).

This method of purifying G protein is particularly useful to isolate G proteins that bind to the Con-218 proteins in the absence of an activating ligand.

Some of the preferred embodiments of the invention described above are outlined below and include, but are not limited to, the following embodiments. As those skilled in the art will appreciate, numerous changes and modifications may be made to the preferred embodiments of the invention without departing from the spirit of the invention. It is intended that all such variations fall within the scope of the invention.

The entire disclosure of each publication cited herein is hereby incorporated by reference.